Title
Agent-Based Modeling of Complex Spatial Systems

Permalink

Authors
National Center for Geographic Information and Analysis, UCSB
Center for Advanced Spatial Analysis, University College London

Publication Date
2007-04-01

Workshop—
Agent-Based Modeling of Complex Spatial Systems

April 14-16, 2007

Santa Barbara, CA

Over the past few years two research communities have developed more-or-less independently: the community of agent-based modelers of spatial systems on the one hand, and the community interested in the representational and computational aspects of complex dynamic systems on the other.

As part of the joint **US National Science Foundation–UK Economic and Social Research Council Special Activity in the Area of E-Science**, the **University of California, Santa Barbara** and **University College London** received funding for a workshop of approximately 30 participants.

The event was hosted by the **National Center for Geographic Information and Analysis (NCGIA)** and UCSB at the Upham Hotel in Santa Barbara, California. The workshop followed the traditional format of a specialist meeting of the **NCGIA**, combining presentations with plenary and small-group discussions.

Presentations

Session I: Representation in and of Complex Spatial Systems

David Bennett, University of Iowa
May Yuan, University of Oklahoma
Discussant: Georgios Theodoropoulos, University of Birmingham

Session II: Modeling Complex Spatial Systems

Mike Worboys, University of Maine
Marina Alberti, University of Washington
Discussant: Nigel Gilbert, University of Surrey

Session III: Validation and Verification of ABMs

David O'Sullivan, University of Auckland and Mark Gahegan, Penn State
Dawn Parker, George Mason University
Discussant: Mike Batty

Agent-Based Modeling of Complex Spatial Systems—Participant List

Marina Alberti
University of Washington
Li An
San Diego State University
Marc Armstrong
University of Iowa
Richard Aspinall
Macaulay Institute
Mike Batty
University College London
Jay Bayne
Meta Command Systems
David Bennett
University of Iowa
Mark Birkin
University of Leeds
Dan Brown
University of Michigan
Matthew Collier
University of Oklahoma
Crispin Cooper
University of Cardiff
Helen Couclelis
UC Santa Barbara
Tom Cova
University of Utah
Catherine Dibble
University of Maryland
Hamid Ekbia
Indiana University
Andy Evans
University of Leeds
Tom Evans
Indiana University
Mark Gahegan
Pennsylvania State University
Nigel Gilbert
University of Surrey
Alan Glennon
UC Santa Barbara
Michael Goodchild
UC Santa Barbara
Richard Harris
University of Bristol
Alison Heppenstall
University of Leeds
Kathleen Hornsby
University of Maine
Andy Hudson Smith
University College London
Indy Hurt
UC Santa Barbara
Kevin Johnston
ESRI
Naicong Li
University of Redlands
Paul Longley
University College London
David Maguire
ESRI
George Malanson
University of Iowa
Steve Manson
University of Minnesota
Harvey Miller
University of Utah
Richard Milton
University College London
David O'Sullivan
Auckland University
Dawn Parker
George Mason University
Donna Peuquet
Pennsylvania State University
Edoardo Pignotti
University of Aberdeen
Rob Procter
University of Manchester
Edward Pultar
University of Utah
Victor Schinazi
University College London
Raja Sengupta
McGill University
Georgios Theodoropoulos
University of Birmingham
Andy Turner
University of Leeds
Stephan Winter
University of Melbourne
Ouri Wolfson
University of Illinois at Chicago
Mike Worboys
University of Maine
Belinda Wu
University of Leeds
May Yuan
University of Oklahoma

Participant Position Papers
start on page 3

Modeling Complex Urban Landscape Dynamics: A Pattern-Oriented Hierarchical Approach

Workshop on Agent-Based Modeling of Complex Spatial Systems

Santa Barbara April 14-16, 2007

Marina Alberti
Department of Urban Design and Planning
University of Washington

As complex dynamic systems, urban landscapes emerge from the local interactions of socioeconomic and biophysical agents and processes. These complex systems are highly heterogeneous, spatially nested, and hierarchically structured (Wu and David 2002). They are prototypical complex adaptive systems, which are open, nonlinear, and highly unpredictable (Levin 1998, Portugali 2000, Gunderson and Holling 2002). Patterns emerge from numerous locally made decisions involving multiple human and biophysical agents interacting among themselves and with their environment. These agents are autonomous, adaptive and change their rules of action based upon new information. Interactions within this complex domain between agents and processes are scale dependent.

While important progress has been made in modeling complex human and natural systems, the ability to simulate emergent behavior in ways that reasonably capture patterns observed in urban landscapes remains a significant research challenge. One major challenge in modeling urban landscape dynamics is in representing explicitly the human and biophysical agents at a level of disaggregation that allows us to explore the mechanisms linking patterns to processes (Portugali 2000). A second challenge in modeling the interactions between human and natural systems is that many factors operate simultaneously at different levels of organization. Additionally, since urban landscapes are spatially heterogeneous, changes in driving forces may be relevant only at certain scales (Turner et al. 1995). Yet our current understanding of the interactions between spatial scales is limited. Simulating the behavior of urban landscapes requires not only an explicit consideration of the temporal and spatial dynamics of these systems, but it also requires identifying the interactions between human and biophysical agents across the different temporal and spatial scales at which various processes operate.

A new Biocomplexity research project at the University of Washington (UW) and Arizona State University (ASU) aims to develop a new framework for modeling the complex coupled human-natural system dynamics of Seattle and Phoenix metropolitan areas.* We propose a pattern-oriented hierarchical approach to model how complex agent-based interactions generate landscape patterns at multiple temporal and spatial scales. We hypothesize that similarly to other ecosystems described by Scheffer et al. (2001), in urban landscapes changes from one state (characterized by a set of processes) to another (characterized by a new set of processes) can be triggered either by the action of slowly changing variables or by relatively discrete shocks. We hypothesize that urban landscapes are spatially nested hierarchies in which the hierarchical levels correspond to structural and functional units (Wu and David 2002). Using a hierarchical modeling approach we aim to identify the structural and functional units at distinct spatial and temporal scales of human and biophysical processes and specify the agents and rates of processes that characterize and distinguish the levels in the hierarchy. The hierarchical patch dynamics perspective emphasizes both the vertical structure (linkages between scales or organizational levels) and horizontal structure (spatial patterns) of the urban landscapes (Wu and David 2002). This perspective allows for a more realistic representation of the relationships among patterns, processes, and scales that lead to emergent properties of heterogeneous urban landscapes.

We specify this model using a spatially explicit, agent-based approach. The model will incorporate the hierarchical patch dynamic modeling approach: such a strategy allows an explicit representation of the nested organizational hierarchies present in human-biophysical systems and thereby provides an elegant means of understanding the interconnections between hierarchical levels. We implement this approach using a hybrid method that integrates dynamic probabilistic relational model (DPRM) and an agent based model. Using DPRM, parameters and spatial rules are estimated empirically from two longitudinal land cover and land use data sets developed for the Seattle and Phoenix Metropolitan Areas.

*For a description of this project see: BE/CNH: Urban Landscape Patterns: Complex Dynamics and Emergent Properties. Alberti, M. (PI), Wu, J., Redman, C., Marzluff, J., Handcock, M. Anderies, J. M., Waddell, P., Fox, D. and H. Kautz. NSF Biocomplexity 2005-2009.

References

Gunderson, L.H., and C.S. Holling. 2002. *Panarchy: Understanding Transformations in Human and Natural Systems.* Island Press, Washington , DC

Levin, S.A. 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems 1: 431–436.

Portugali, J. 2000. Self-Organization and the City. Springer, Berlin.

Scheffer, M., S. Carpenter, J. Foley, C. Folke, B. Walker. 2001. Catastrophic Shifts in Ecosystems, Nature 413: 591-596.

Wu, J. and David, J. 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecological Modeling, 153: 7-26.

Personal perspectives on ABM:

Development of the Pseudo-history Approach

Li An

Questions: Sampling is a fundamental way to collect data for scientific investigations in various disciplines. For instance, scientists have been employing sampling data to test hypotheses, investigate new phenomena, and establish new theories. However, people have seldom, if ever, asked questions regarding the geographic size of the sampling frame, the time span of data collection (for longitudinal studies), and the frequency of data collection. Decisions related to these questions are usually made based on the researcher's experiences, data availability, or "common sense" without exploring whether the relationships under investigation can be captured by data collected using the chosen sampling strategy. An example in Geography is that researchers choose spatial data sets (e.g., satellite images) of their study site over a certain time span and at a certain time interval, which are largely decided by factors such as costs and availability of the data. This lack of examining the match between the extent, scale, and rate of the real processes (often not directly observable) under investigation and the corresponding sample strategy may substantially undermine the validity of the subsequent analysis and findings. Many reasons may account for this problem, such as <u>uncertainties</u> in the processes and the related system structures of interests, difficulties in observing and quantifying the spatial and temporal scales of some phenomena, and <u>costs</u> of data collection in some large areas, over long time, and at short time interval.

Justification: The advent of fast computers in the last decades and advancement in software have brought forward new opportunities to address the above sampling strategy questions using agent-based spatial models. Rather than solely depending on real data that are sometimes subject to the above constraints, researchers can use computers to generate an artificial digital "world", and let whatever processes of scientific interests proceed on this 'world". Theories, experiences, and hypotheses could be used to guide or affect the directions, strengths, and interactions of such processes. The emergent patterns on the artificial world, sometimes unexpected and surprising, may arise from the behaviors of many agents (autonomous entities or objects that have goals, some degree of knowledge, and actions), the interactions among themselves, and their relationships with the environment(s). This type of "realistic" <u>mapping</u> of real-world entities and processes onto a computer model facilitate a totally new scientific investigation in comparison with traditional analyses based on empirical data: computational simulation. A digital artificial laboratory can be constructed to take major relevant entities and processes under certain research objectives and eliminate unimportant details and noises, thus reducing uncertainties, difficulties, costs, which would otherwise prohibit many data collection, analysis, and scientific investigation activities. Thus Computational simulations undoubtedly will bring a new era in addressing the above questions relating to the match between the extent, scale, and rate of real processes (but often not directly observable) and our sampling strategy.

Because the researcher controls the processes that underlie the observable patterns through setting the parameters or choosing the algorithms, we can freely choose varying sampling strategies by setting the simulation time frame, automating sampling at the designated intervals, and choosing a certain geographic area within our artificial world. Analyzing data obtained through such strategies with the knowledge of the processes ongoing should be able to reveal the effectiveness of different sampling strategies. However, such a crucially important area has seldom, if ever, been touched. For instance, a web of science search under this combination of key words "(sample or sampling) and (artificial world or computational simulation)" resulted in two irrelevant records (this search is not exclusive, but may be still suggestive).

The pseudo-history ABM approach: A new approach developed by An and Brown (in preparation); see CV, the pseudo-history analysis, can be used as a socio-environmental laboratory that accommodates all the sampling strategy related needs, conduct computer-based experiments, and advance scientific research related to the effectiveness of sampling strategy. The current pseudo-history model encapsulates a hypothetical "landscape" that is available for human residence and service center development at the beginning. The "landscape" can vary by size, and several variables (parameters) control its environmental and socioeconomic features such as soil quality and distance to service center. As time goes on, i.e., as computer internal clock moves, "homebuyers" (computer objects) with various characteristics and preferences (parameters) enter the "landscape", evaluate different locations (controlled by algorithms and equations) as candidate locations, and choose locations that meet their objectives of maximizing residence utility.

Given such a hypothetical "landscape" and the residence-choice related processes on it, the following tasks can be implemented in order to achieve the above objectives: (1) further verify the current model, including debugging, (2) let some residence decision-related parameters correlat with changes in the environmental or socioeconomic parameters, (3) add new features to the model, such as letting the residence-choice and land-use processes have higher correlation with time and space, automating the model in taking samples at varying sizes, spans, and time intervals, and writing the sample results in appropriate formats, and finally (4) examine what sampling strategies in relation to geographic sampling frame, sampling span, and sampling interval would best disclose model inputs that have generated the sample data. Furthermore, because the researcher knows the true mechanisms ongoing and their temporal/spatial scales (established as parameters and rules in the ABM), and thus many uncertainties are removed, the pseudo-history approach can be used to test what statistical methods can best detect the underlying rules accommodated in the ABM.

Significance: This approach is significant for several reasons. First, it explores the effectiveness of varying sampling strategies, which would benefit scientists who use sampled data regardless of what fields they come from. Second, it will advance methods and theories of the Geocomplexity, an increasingly recognized field that connects Geography, computer science and engineering, and complexity science, witnessing tremendous applications in many theoretical and empirical studies. Last, this approach will add crucial components to a spatial modeling and simulation curriculum, and give graduate students invaluable experiences in computational modeling and simulation.

NSF/ESRC Agenda Setting Workshop on
Agent-based Modelling of Complex Spatial Systems

Richard Aspinall
Macaulay Institute,
Craigiebuckler, Aberdeen, AB15 8QH, Scotland, UK
r.aspinall@macaulay.ac.uk

Land-use change as an application that challenges capabilities in agent-based modelling and spatial simulation

Modelling and analytical needs of the science land-use and land-cover change present some demanding challenges to agent-based and spatial modelling methodologies. Land-use change is increasingly recognised as an emergent property of interactions of coupled natural and human systems operating as a complex adaptive system (Stafford-Smith and Reynolds, 2002; Lambin *et al*, 2003). Land-use and land-cover change have been related to a variety of direct (proximate) and indirect (underlying) factors[1] (Geist and Lambin 2002, 2004). Observed changes are associated with i) multi-factor explanations, including interaction between factors, ii) complex local- and regional-scale institutional and individual decisions, themselves related to national- to global-scale opportunities associated with new technologies as well economic and other policies, and iii) historical contingency, reflecting development and transition of underlying and proximate factors over time and producing both path dependence and non-stationarity in change (Aspinall, 2004). Additionally, the meta-analysis of case studies of land-use change by Geist and Lambin (2002, 2004) shows that no universally applicable (in space or time) policies or practices for policy-level direction of land-use change are found, and that a detailed understanding of change at a given location is required to evaluate place- and time- specific patterns of change.

Explanation and modelling of change based on driving factors has served reasonably well, both as a basis for comparison across an international suite of case studies and for identifying important sets of influences on land-use change. However, for improved understanding and ability to analyse and model land-use change a factor-based approach must be expanded to include explicit recognition of processes that produce change. Natural system processes influencing land-use change include a set of processes concerned with soil, climate, ecosystems, and hydrology and there has been considerable progress in coupling environmental process models to GIS for scientific study of natural systems at spatial scales from local to global and time scales from short to long. The human system processes influencing land-use change reflect a complex set of individual, group, and institutional decision-making; there are fewer models of these processes despite efforts in economics, and an increase in agent-based modelling as a mechanism for addressing decision-making processes.

These qualities of systems of land-use and land-cover change indicate some of issues that agent-based and spatial models need to be able to address to i) improve understanding of processes producing change, ii) model change, and iii) improve our

[1] Underlying factors include demographic, economic, technological, policy and institutional, cultural factors (Geist and Lambin, 2002, 2004)

ability to predict change and project into the future. Two general requirements relate to the observation that there are no universally applicable policies or practices for direction of land-use change:

1. Place-based analysis and models are necessary for understanding land-use change. The capabilities of GIS to support place-based analysis and modelling provide a foundation for this.

2. Time-based (including historical context and contingency) analysis and modelling are necessary for understanding land-use change. Modelling methodologies and technologies thus need to be able to include characteristics of temporal context and contingency explicitly. This complements the need for place-based analysis to provide models that can adequately reveal local responses to change.

In the context of agent-based models that address individual, group and institutional issues, including decision-making processes, which operate primarily within the set of factors associated with the human systems component of land-use change I identify three related needs:

1. Agent-based models are needed that can represent individual, group, and institutional decision-making, and their interaction, in the context of larger scale economic and other (e.g. technological) opportunities and policies.

2. Management of the spatial representation of landscapes linked to an agent-based model such that the spatial extent, behaviour, and actions of individual agents are adequately located in space. This includes using GIS to provide spatial context and identify neighbours for 'geographic agents' in order to provide potential for peer-group interactions, diffusion of innovation, and shared or common decisions or responses that influence land-use change.

3. Management of the temporal resolution of agent-based models in relation to the temporal application and behaviour of the entity being represented as an agent. For example, many economic policies or technological opportunities effectively operate as constants over some period of time, yet individual, group and institutional responses to these policies and opportunities evolve and otherwise change over time.

Aspinall, RJ (2004) Modelling land use change with generalized linear and generalized additive models – a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. *Journal of Environmental Management*, 72, 91-103

Geist, HJ and Lambin, EF (2002) Proximate causes and underlying driving forces of tropical deforestation. *BioScience*, 52, 143-150

Geist, HJ and Lambin, EF (2004) Dyanamic causal patterns of desertification. *BioScience*, 54, 817-829.

Lambin, EF, Geist, HJ, and Lepers, E (2003) Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environmental Resources*, 28, 205-241.

Stafford-Smith, DM and Reynolds, JF (2002) Desertification: a new paradigm for an old problem. In: *Global Desertification: Do Humans Cause Deserts?* Reynolds, JF and Stafford Smith, DM (Eds), Berlin: Dahlem University Press, 387-401.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Evaluating Agent-Based Spatial Models

Michael Batty, CASA, UCL

I am interested but worried about the development of this relatively new class of models which tend to fight against very long standing principles in science that seek for simplicity in abstraction and application. Agent-based models (ABM) have developed for spatial systems through advances in computing that enable distinct objects to be defined with respect to their behaviours which in turn suggest that much more highly disaggregate systems can be represented than hitherto. The fact that ABMs are defined with respect to behaviour implicitly means that such models are temporally dynamic and this too increases the richness of these models in terms of the variability of the systems that are being simulated. Spatial data too is being acquired at ever finer spatial and temporal scales and is making possible the development of richer models structures of which ABMs are typical.

I consider that the workshop should spend time discussing various types of ABMs, defining different types at different levels of object specification, scale, and spatio-temporal disaggregation. Not all ABMs are equated with large scale data requirements that are hard to meet and thus there may be model types that do not conflict with the hallowed canons of parsimony that have defined the scientific method and the model-building process hitherto. Moreover we need to consider ABMs that are not designed to mirror a reality in terms of religiously simulating an observed system. Some models are designed for much more general purposes in terms of defining baselines, structuring data, and enabling hypothesis generation, and thus we need to be clear about the purpose for which particular models are being built.

My own view is that most ABMs are being developed for purposes that are similar to those for which models in the social science have traditionally been constructed: for making predictions that inform policy in the future, although there is a distinct subgroup of such models that are focussed on the past and these tend to break this symmetry[1]. In fact, where ABMs do not work very well in terms of their generating predictions that are close to the reality they seek to simulate, the purpose for which they are initially defined tends to change to less ambitious aims. What has clearly happened however is that as new types of models such as ABMs have emerged over the last 20 years, the whole process of confronting models with reality through observational data, has been elaborated. Here we will refer to the process of matching the model against reality and theory as ‘evaluation’ which will include different aspects of this confrontation which are called <u>verification</u>, <u>validation</u>, <u>calibration</u>, <u>confirmation</u>, <u>falsification</u>, <u>prediction</u>, and <u>accreditation</u>.

[1] van der Leeuw, S. E. (2005) Why model? **Cybernetics and Systems**, **35**, 117-128.

Accreditation we can probably dispense with as it involves the extent to which the model comes from credible sources, although there is always the prospect of models emerging from outside the scientific establishment.

If a model is designed to predict some reality, then the only unique test that is acceptable is one in which the model's predictions are compared with observations that are entirely independent of the information used in constructing the model in the first place[2]. If a model is fitted to an existing set of data which contains variables that are defined as inputs and outputs with the model being based on associations between these inputs and outputs, the fact that the model predicts the outputs correctly from the inputs is not an acceptable test of the model's quality as the data used is not independent. The only acceptable test would be for the model to be used on another set of data which is judged to be independent of the data set in question. In terms of spatio-temporal models, this data set should be at a different time in a different space. Even models that are fitted to a space at time t and generate predictions for the same space at time $t+1$ are unacceptable for the data is not completely independent. In fact in spatio-temporal systems, this is an extremely hard criterion to follow and it can damn entire classes of model in terms of determining their predictive abilities. Although this is often confused with verifying a model, we call this process of testing a model's predictions against an independent set of observations validation. The process of fitting a model to the data set around which is has been designed is in fact called calibration which is akin to statistical estimation in that this is the way the unknown parameters are determined which fine tune the model to a real situation. In fact, the process of validation and calibration might be one and the same in that what this implies is that at least two sets of data are needed, one on which to estimate the model and one on which to gauge its predictions from these estimations. Sometimes a single data set is partitioned into two, with estimation occurring on one and validation on the other. Sometimes it is said that the model is trained to the first subset so that when trained it can be used to predict the second subset of data.

The process of verification is defined here as one in which the model's assumptions and construction is checked for consistency and plausibility. A good definition by Miser and Quade[3] is "…the process by which the analyst assures himself and others that the actual model constructed is indeed the one he intended to build". They also contrast with the definitions of validation as "…the process by which the analyst assures himself and others that the model is a representation of the phenomenon being modelled that is adequate for the purposes of the study…".In short, verification consists in generating some sense in which the model is consistent with theory, produces logical results, and is

[2] Manson, S. M. (2007) Challenges in Evaluating Models of Geographic Complexity, **Environment and Planning B**, **34**; forthcoming. http://www.envplan.com/contents.cgi?journal=B&volume=forthcoming

[3] Miser, H. J. and Quade, E. S. (1988) Validation, In Hugh J. Miser and Edward S. Quade (Editors) **Handbook of Systems Analysis: Craft Issues and Procedural Choices**, North Holland, New York, 527-565; quoted in Hodges, J. S. (1991) Six (or So) Things You Can Do with a Bad Model, **Operations Research**, **39**, 355-365.

appropriate to the problem being explained. These issues also appear during validation but it is possible to verify a model without validating it and vice versa.

In validation which involves comparing model outputs and independent observations, model predictions might be confirmed or falsified. In fact, this process is always ambiguous for there is always uncertainty in the data which confronts the model, and there is always uncertainty in calibration. In short, predictions are never perfect and thus there are important value judgements to be made in the process of validation. For ABMs, the balance of calibration, verification and validation needs to be explored in considerable detail for models which lack parsimony as most ABMs do, contain assumptions and processes that cannot be compared against data, and thus the balance between verification and validation is different from those model processes and outputs which an be so compared with data. In discussion, I hope we will introduce various examples that illustrate these difficulties. From my own work and some of my associates, I will show how models of local movement where actual behaviour can be hypothesised quite plausibly but rarely observed in terms of individual movements defy full validation, and require less than best strategies to be designed for their evaluation.[4]

Many of the models that we will discuss at this workshop cannot be validated in the traditional sense and this requires us to be very clear as to the purposes for which they are constructed. Hodges and Dewar[5] define several distinct reasons for building models that cannot be validated and these revolve around notions of using model for 'book-keeping', for selling an idea, for training, for management, for communication, for strengthening arguments, and for generating new theory and hypotheses. I am particularly intrigued by the class of agent based models that deal with disaster management such as those dealing with crowding where the models are potentially near to full validation but the circumstances of such validation may not be repeatable – i.e. disasters – and are certainly not ethically desirable to be repeated. These events are also conditioned by extreme control and this makes their observation unstable through time. Helbing's work[6] on the disasters at the Hajj, the Pilgrimage to Makkah, count amongst these examples. I would be fascinated in what the balance of opinion on these questions is amongst our expert group as I imagine that this will depend as much in disciplinary perspective and on the extent to which we consider our models as being central to policy-making.

---

[4] Batty, M., Desyllas, J., and Duxbury, E. (2003) Safety in Numbers? Modelling Crowds and Designing Control for the Notting Hill Carnival, **Urban Studies**, **40**, 1573-1590.
[5] Hodges, J. S., and Dewar, J. A. (1992) Is It You or Your Model Talking? A Framework for Model Validation, RAND Corporation, **R-4114**, Santa Monica, CA.
[6] Helbing, D., Johansson, A., and Al-Abideen, H. Z. (2007) The Dynamics of Crowd Disasters; http://arxiv.org/pdf/physics/0701203; and http://www.trafficforum.org/crowdturbulence

Issues of Representation and Interpretation for Agent-based Models of Complex Adaptive Spatial Systems

David Bennett, Department of Geography, The University of Iowa

What we know of the world around us is, in large measure, the product of reductionist science. The basic tenets of this approach tell us that truth can be found through an understanding of individual system components; a system is the sum of its parts. While this approach served us well through much of the 20th century, many scientists now believe that a reductionist approach alone is insufficient for the study of natural and social systems. These scientists promote a new approach based on complexity theory and complex adaptive systems (CAS) that is focused as much on the linkages among system components as the components themselves; a science where the underlying assumption is that a system can be more than the some of its parts. Traditional scientific methods, however, are often ill-suited to the study of complex systems characterized by feedbacks and non-linear dynamics, path dependency, adaptation, cross-scale interaction, self-organization, emergent behavior, and dissipative processes. Agent-based modeling (ABM) has been highly touted as an appropriate technique for the study of complex adaptive spatial systems (CASS).

My interests in ABM for CASS lie primarily in the representation of intelligent, mobile, spatially-aware, and adaptive decision-makers. More specifically, my research has been focused on how individuals make decisions about: 1) land use, land cover, and associated management strategies; 2) how to navigate across uncertain and risky landscapes (elk in this situation); and 3) how to organize to effect change in policies that, in turn, effect changes in the production of ecosystem services. Linking all three of these projects are underlying questions about how landscape structure emerges from individual and localized action and how feedback mechanisms link multiple social or spatial scales. Gaining an understanding of landscape-scale processes through the use of ABM presents significant challenges for the representation of spatially-aware cognitive agents and in the interpretation of model results. These challenges must, in my opinion, be addressed before ABM will meet our high expectations for the study of CASS. In the following discussion I lay out some of these challenges in greater detail.

Representational challenges

Two related and significant challenges for the development of ABM for CASS are the representation of cognition and context. Complexity is often discussed in terms of self-organization and emergent behavior; behavior driven, in part, by adaptive processes. For humans (and presumably other higher order animals), short-term adaptation requires cognition. Research is needed on how to represent and implement cognitive processes (learning, reasoning, and memory) in ABM for CASS. For example, spatial decision-making is often a collaborative, multi-objective, and semi-structured process supported by limited and uncertain knowledge. Agents built to support the simulation of CASS might, therefore, be required to learn to: 1) manage spatial resources under uncertainty; 2) organize, compromise, and collaborate to reach individual or societal objectives; and 3) minimize risk and maximize opportunity. “Hard-coding” these behaviors into a system is likely to lead to what Holland (1986) has referred to as “brittleness” and a failure to capture complex behavior. Unique to the simulation of spatial systems is the need to represent spatial cognition. Learning safe routes through a landscape may require, for example, the digital equivalent of cognitive maps that agents learn, store, reference, and adapt to changing risk surfaces. While there has been a significant amount of work done in machine learning for more simplified environments (e.g., robotics), little of this kind of work has found its way into models of complex adaptive spatial systems.

Cognitive behavior is, generally speaking, derived from a history of contextualized experiences. Spatially-aware, intelligent agents must, therefore, connect external stimuli (e.g., resources and threats), internal states (e.g., wealth, nutrition, social connections), and the states and behaviors of other agents (neighbors, colleagues, competitors) to successful behavior and generalize this knowledge to similar situations. Furthermore, an appropriate spatial response might depend on a particular sequence of events.

Context given heterogeneous agents with bounded knowledge about complex spatial systems must, therefore, be agent-specific and derived from a spatio-temporal representation.

Interpretational challenges

When we use simulation we need to know that the model accurately reflects the real-world processes of interest and that this model was accurately translated into software. The goal of complex system models is often to explore system-level behavior as it is produced by a large number of interacting and heterogeneous agents. ABM are often large and complicated, which makes model verification and validation challenging. Much has been written about the verification and validation of ABM and, while it remains a significant challenge, I will make just two quick comments here. The first comment (an admittedly obvious one) is that the identification and resolution of verification and validation problems becomes even more difficult given a virtual system expected to produce complex non-linear, stochastic, path dependent, and emergent behavior. If we accept, for example, an explanation based on complexity then we must also accept that an existing spatial pattern is just one realization of many possible alternative states. A model that fails to reproduce the existing state is not, therefore, necessarily in error. Similarly, the concept of equifinality suggests that a model that does mimic real-world patterns is not necessarily valid (Brown et al. 2006). Second, the use of ABM in CASS makes most sense when one is studying how the actions and interactions of individuals lead to system-level behavior. Relations among individuals or between individual and their environment at a single analytical scale can often be studied more directly using other techniques (e.g., a statistical approach). However, models of adaptive, contextually aware agents are complicated and the output difficult to interpret. How do we prove, for example, that the emergent behavior (ignoring for now how this is defined and measured) produced by the system is, in fact, generative evidence of real complex behavior, and not an unintended artifact of some simplifying assumption encoded into agent behavior?

This brings me to the final issue that I wish to address in this position paper. If a generative scientific approach, like ABM, is to be applied to CASS it must be transparent. We might expect system dynamics to be transparent simply because agent behavior is explicitly encoded, but issues of adaptation, equifinality, bifurcation, and divergence, the very behaviors we expect the system to capture, can quickly render the modeling process opaque. It makes sense to build into complex system models the same kinds of explanatory tools typically associated with expert systems, but tracking cause and effect for a CASS will be considerably more complicated. Can we determine *a priori* what an important event in an ABM simulation looks like? When, for example, is the variation in the state of some modeled component unimportant noise and when does it signal a bifurcation point? Building into ABM the ability to trace back through model output to gain an understanding of how a system got to where it did is likely to prove challenging, but it seems imperative that we do so if we are to make strong claims about our interpretations of model results. The first step toward such a capability might be the construction of the kinds of cognitive and spatio-temporal data representations discussed above.

Final thoughts

Given the challenges associated with cognition, context, and model interpretation, what conclusions can be drawn from ABM about complex spatial systems? A goal of statistically valid models of real world processes seems a ways off and, perhaps, even misguided. Perhaps the greatest value of agent-based models for complex adaptive spatial systems lies in the questions that they require us to ask about system behavior, the way that they require us to conceptualize system structure, and the opportunities that they provide for us to explore plausible outcomes and search for robust decisions.

Brown D.G., Aspinall, R., Bennett, D.A. 2006. Landscape Models and Explanation in Landscape Ecology—A Space for Generative Landscape Science? *The Professional Geographer*, 58(4): 369–38.

Holland. J.H. 1986. Escaping brittleness: The possibility of general-purpose learning algorithms applied to parallel rule-based systems, In: *Machine Learning, an Artificial Intelligence Approach, Morgan-Kauffman, vol. 2*.

Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

Agent-Based Modeling of Complex Spatial Systems

UCSB April 14-16, 2007

Jay Bayne, PhD
jbayne@metacomsys.com

We are interested in the physics of *cyberspatial objects* (events, agents, processes) whose behaviors unfold in *cyberspace*, a manifold where geospatial, infospatial and temporal indices are required to distinguish one object from another and, for a given object, one state from another. Consequently, we propose a c*yberspace-time* (CST) reference framework for describing the dynamics of *cyberphysical systems* (CPS). A CPS is an information system (object, intelligent agent or event) whose behavior is defined in geospatial, infospatial and temporal terms. A CPS is responsible, in whole or in part, for observing and possibly reacting to cyberphysical events. As such it may be an observer-controller of other CPS, in client-server (peer-peer) or parent-child (superior-subordinate) configurations. CST is the cyberspatial analog of the classical space-time reference frames found in Newtonian and relativistic physics.

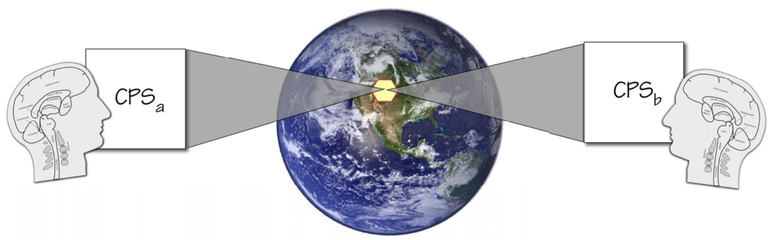


Figure 1 – Situation Assessment

Figure 1 expresses graphically the following problem: Two autonomous observers (rational[1] cyberphysical systems CPS_a and CPS_b), from their respective independent reference frames, are tasked with monitoring events related to a specific geophysical process. How do they identify, discover, bind to and observe (measure) the same process? How do (should) physical processes and their proxies (e.g., sensor and actuator subnets) represent their states and behaviors? How do the independent CPS observations compare; by what metrics? What differences arise in their perception of events when viewed from distinct cyberspatial reference frames? What are the sources of these variations? How are these variations rationalized? How are events to be time stamped, sequenced, archived and reported (shared)?

These questions, and many others, lead to requirements for a unified *cyberspace-time framework*, one that relates geospatial (physical), infospatial (logical) and temporal dimensions.

Geospatial References

The geospatial reference frame is the familiar three-dimensional Newtonian space-time framework (Figure 2), adjusted when speeds dictate, for relativistic effects by the Loretnz transformation. We are concerned with cyberphysical systems operating within the Earth's biosphere (Gaia), a region roughly 9 km above and below the Earth's mean sea level. Within this region a CPS is tasked with locating, identifying, tracking, reporting on and possibly controlling *cyberspatial objects* (CSO). Naturally, we require a unified means of logging (cataloging) the characteristics and behavior of CSO. An appropriate repository (database) in turn requires a geospatial schema competent to capture both static and dynamic attributes of the enormous population of potential CSO, individually and in various combinations. We therefore require a digital representation of geospace and its contents, a requirement addressed by a *digital earth reference model* (DERM, ref. Figure 3)[2].

[1] An object is *rational* to the degree it chooses actions that are in its own best interests, given the beliefs it holds about the world, its own capabilities, and the constraints under which it is allowed to operate.
[2] We require a bidirectional conversion between classical Euclidean (Cartesian and Spherical) coordinates and DERM coordinates

META COMMAND SYSTEMS ™
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

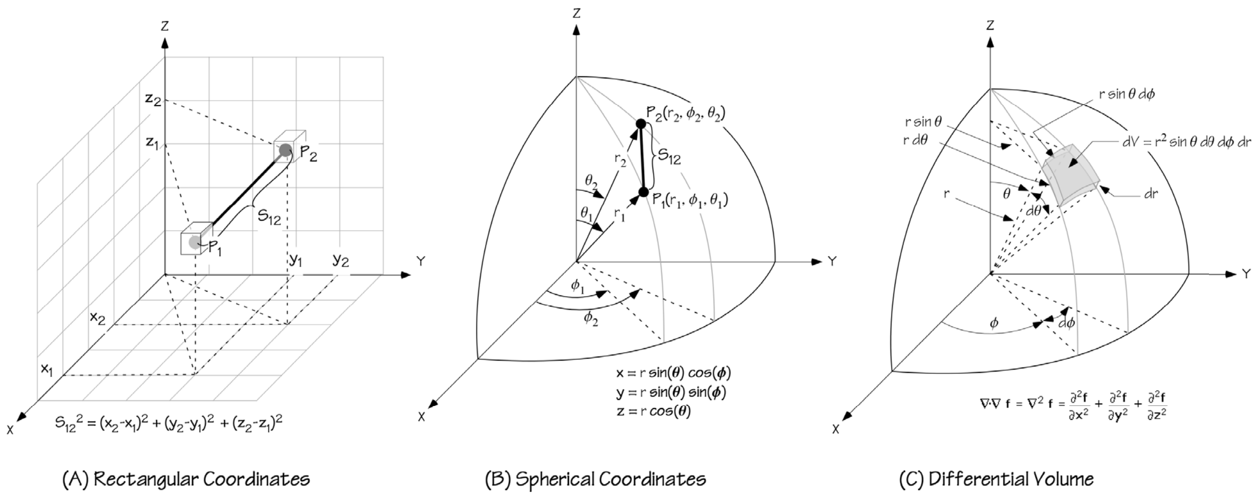


Figure 2 – Geospatial Reference Frames

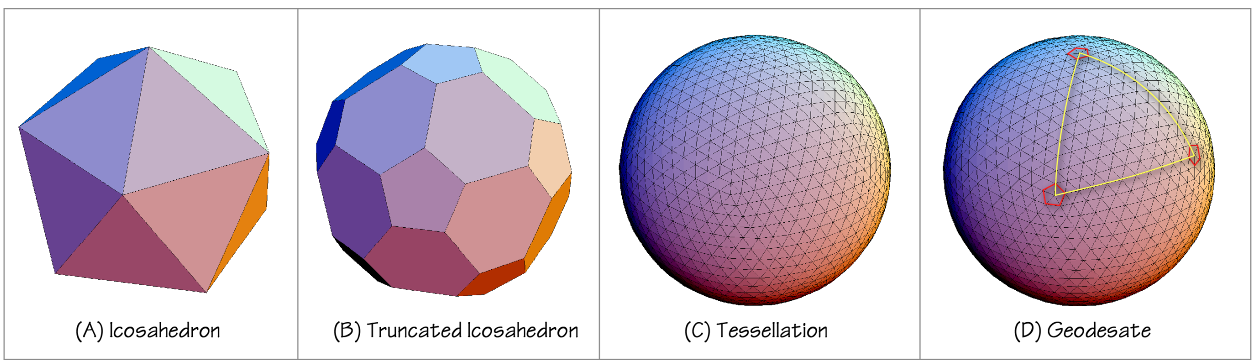


Figure 3 – Aperture 3 Hexagon Tessellations on a Sphere

For reasons beyond the scope of this paper, we chose a particular DERM, one based on the *Pyxis innovation*[3], multiple groups of aperture-3 tessellations placed on an icosahedron and Snyder-projected onto a sphere (Figure 3C). Each subsequent (higher) resolution is related to the previous according to $radius_n = radius_{n-1} / \sqrt{3}$, with alternate tessellations rotated by 30°. Resolution 0, the coarsest, is based on the 12 vertices of the unfolded icosahedron (Figure 3A and Figure 4A). Each icosahedal vertex defines the center of a pentagon (Figure 3D and Figure 4A).

The *Pyxis innovation* indexing syntax, beginning at Resolution 1, is "AN:N…N", where "AN" is an alphanumeric string identifying a vertex or face value, as shown in Figure 4B, and "N…N" is a numeric string giving the standard Pyxis resolution index. In general, a *geospatial address* (*gsa*), expressed in terms of the Pyxis digital earth index, is given by

$$gsa = < dga : dra : dea >$$

The Pyxis DERM Global Address (dga) specifies the Resolution 1 index "AN." The Pyxis DERM Resolution Address (dra) specifies the higher resolution (>1) indices "N…N." And the Pyxis DERM Elevation Address (dea) specifies the thickness (volume) of the cell identified by $< dga : dra >$.

[3] The Pyxis innovation, www.pyxisinnovation.com

META COMMAND SYSTEMS ™
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

Figure 6A emphasizes our interest in events taking place within the Earth's biosphere (Gaia), a region roughly 9 km above and below the Earth's surface. Figure 6B shows the three components of a *Pyxis innovation* geospatial address of a cell at a particular resolution within a hexagonal cone.

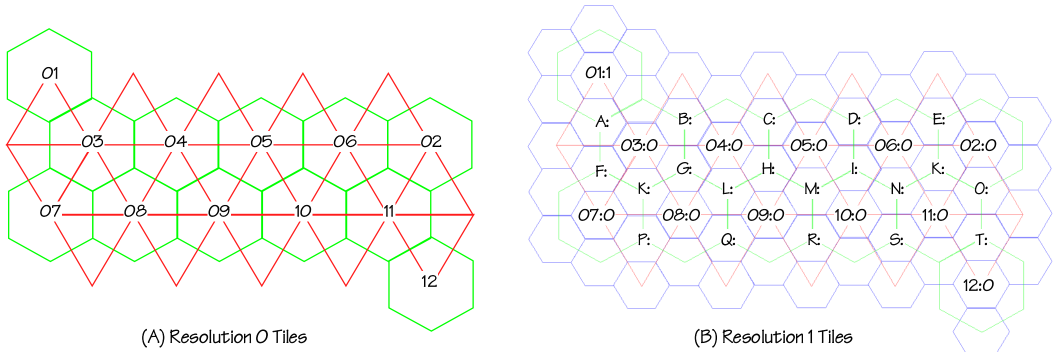


Figure 4 – DERM Indexing

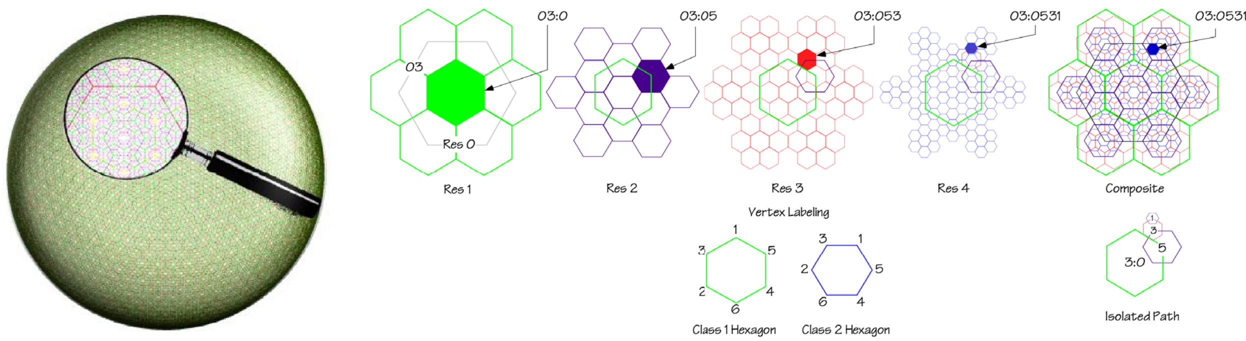


Figure 5 - Nested √3 Hexagonal Tiling (< *dga* : *dra* >)

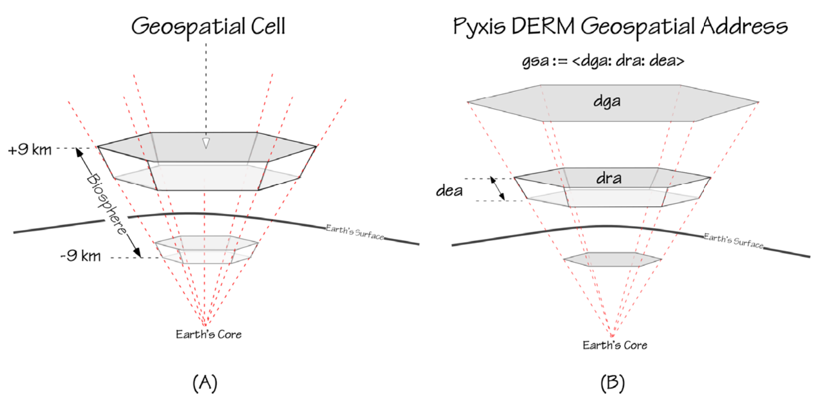


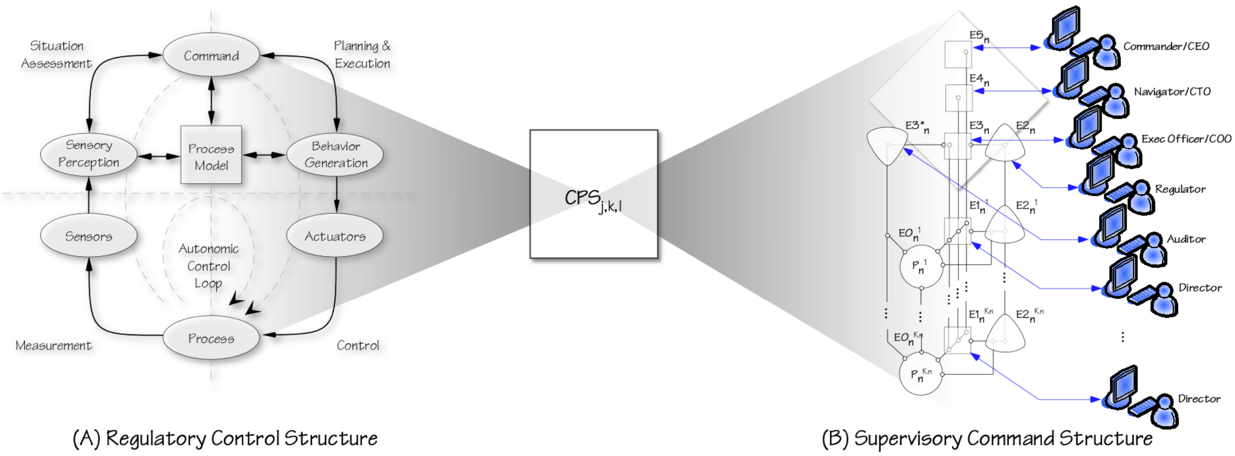
Figure 6 – Identifying Geospatial Cells (Regions with Volume)

Infospatial References

Cyberspatial objects are either static (inert) or dynamic (active). Dynamic objects are *enterprising*, able to communicate, exhibiting both state and behavior. Static or dynamic objects may be of interest to other dynamic objects, independent observers (agents) referred to as *cyberphysical systems* (CPS). As diagrammed in Figure 7, CPS objects are defined in terms of two complementary functions: (A) the processes (*value-added services*) through which they interact (trade) with other objects—their regulatory (production) control structure—and (B) the internal

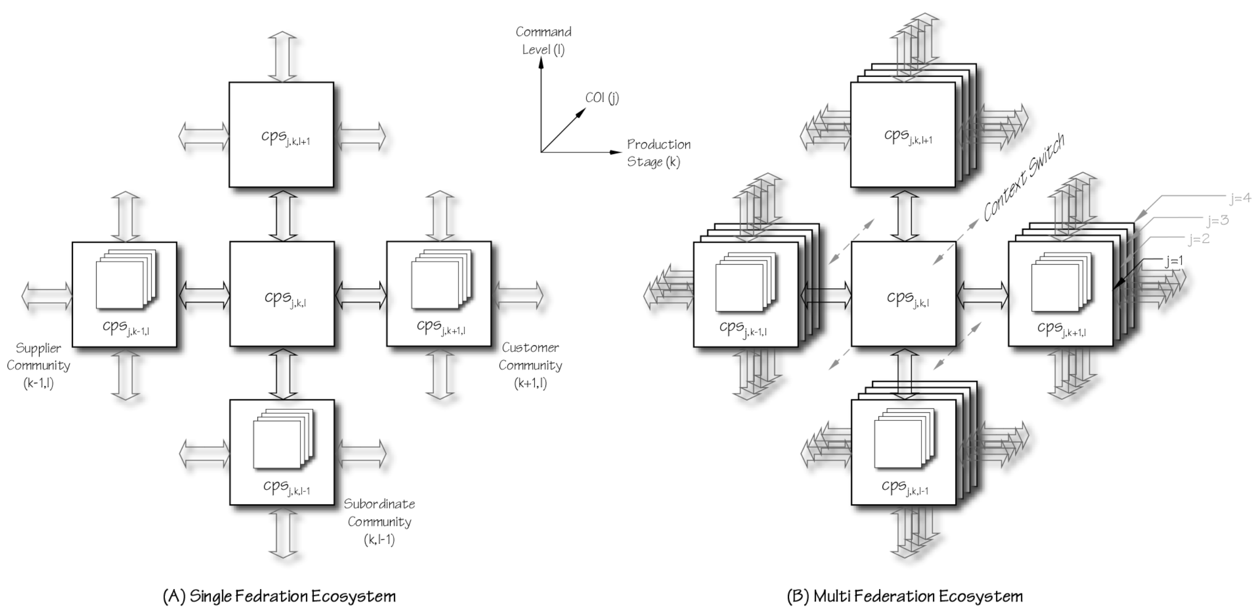
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

governance structure through which they regulate their behavior in order to maintain their viability—their supervisory command (accountability) structure[4].



(A) Regulatory Control Structure (B) Supervisory Command Structure

Figure 7 - A Cyberphysical [Agent] System (CPS)



(A) Single Fedration Ecosystem (B) Multi Federation Ecosystem

Figure 8 - CPS [Agent] Federations

CPS operate in federations (communities if mutual interest, alliances, ecosystems), requiring that each member possess unique identity, a means of discovery and the ability to bind with other CPS for the exchange of goods and services. As described in Figure 8, a CPS operates in a three-dimensional space, within one or more communities of interest (COI), along that community's horizontal production (supply) chain and along its vertical command (asset) chain. Therefore, its identity requires at least three components, one for each axis. In the figure, these components are identified by indices $< j, k, l >$, respectively.

$$< j, k, l > = < federationID, prodIndex, cmdLevel >$$

[4] A detailed discussion of the command and control (C2) structures diagrammed here is contained in several papers available at http://www.metacomsys.com and in the author's text *Creating Rational Organizations – Theory of Enterprise Command and Control*, available at http://www.cafepress.com/mcsi.

META COMMAND SYSTEMS ™
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

$CPS_{j,k,l}$ represents the object's name. In addition to its name a CPS requires unique addresses for its *service access points* (SAP)—ports on the object through which it offers its capabilities (value propositions) to other objects. For identification of service access points we utilize the addressing mechanisms defined in the Internet Protocol version 6 (IPv6) standard.

Figure 8 diagrams the relation among governance structures of members of a single federated system (A) and a multi-federation system (B). This lattice structure emphasizes roles among CPS members along their respective horizontal production and vertical command axes. The multi-federation system (B) also emphasizes the requirement for a CPS *operating system* (CPOS) that supports agile movement (i.e., *context switching, multi-tasking*) among the demands of multiple federated systems.

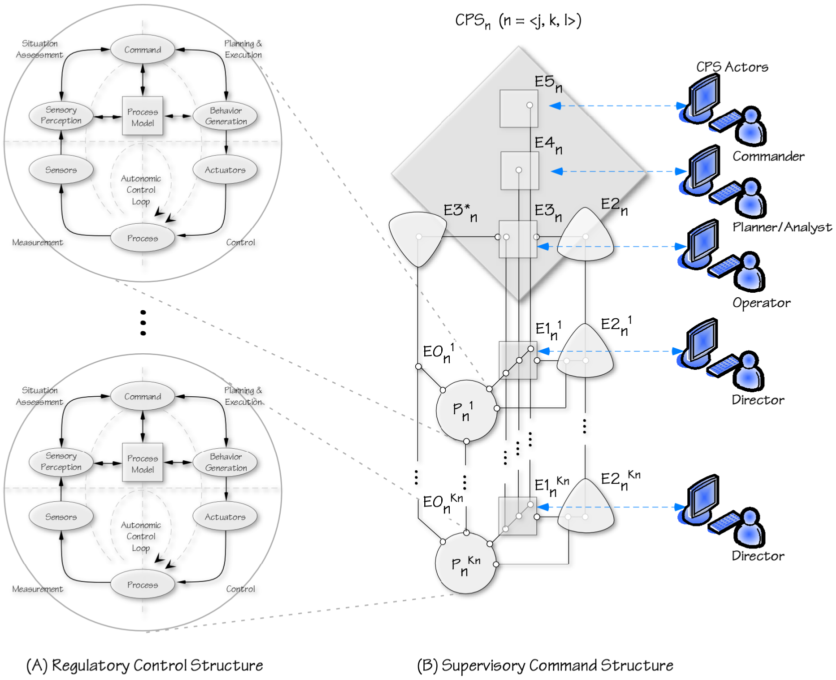


Figure 9 – Replicated CPS Governance Structure

Figure 9 provides a closer look at the governance structure of CPS_n, where $n = f(j,k,l)$. The CPS may contain one or more subordinate *value production processes* $\{P_n^i, i = 1..K_n\}$, each requiring one or more service access point addresses. Typically, a CPS will represent a local subnetwork address space. Figure 10 represents a CPS structure for an enterprise that chooses to distinguish (separate) their intelligence, surveillance and analysis services from their command (decision) and control services. A wide variety of such partitioned service portfolios are possible.

Each P_n^i contains (represents) one or more services of a regulatory control loop. The set $\{P_n^i\}$ are governed (regulated, synchronized) by E_5-E_4-E_3 command structure elements through its two juxtaposed (counter-balancing) feedback loops—the *sympathetic* (E_3-E_2-E_1-E_3) and *parasympathetic* (E_3-E_3^*-E_0-E_1-E_3) aspects of the CPS' autonomic nervous system (ANS). Each E_1 Director is the E_5-E_4-E_3 command structure (rotated 45°) for the next lower (subordinate) level of command. Consequently, each E_0 (P_n^i) represents one or more embedded CPS. This counter-balanced and recursive formulation of CPS governance supports a high degree of scalability and the design and deployment of reusable service-oriented CPS governance software.

META COMMAND SYSTEMS ™
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

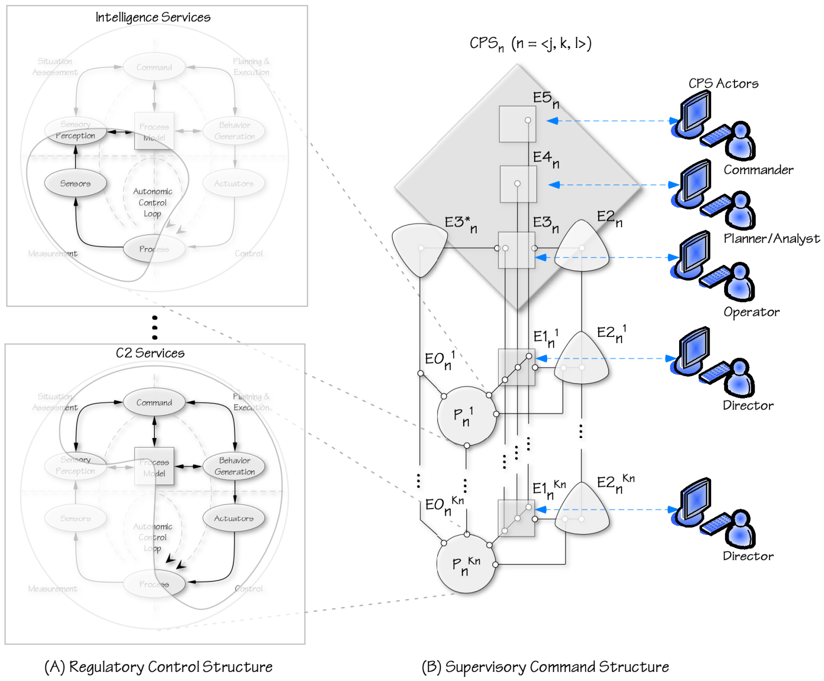


Figure 10 – Partitioned CPS Governance Services

An infospatial reference frame is one that provides a unique naming and addressing (indexing) scheme for cyberspatial system objects that operate within a specific *infosphere*. For our purposes, the highest level "containment domain" is an infosphere expressed in terms of addressing conventions defined in Internet Protocol, version 6 (IPv6) [5]. Within this domain, a communicating object is represented by one or more IPv6 infospatial addresses (ISA).

An ISA is a 128-bit integer identifying one of potentially ~3.4×10^{38} objects. A given ISA may have various formats and interpretations. The general format for an IPv6 address provides the first 48 bits for specifying one of ~2.8×10^{14} Global Network Addresses (GNA), the next 16 bits specifying one of 65,536 Sub Network Addresses (SNA) and the final 64 bits identifying one of ~1.8×10^{19} possible Service Access Points (SAP) within the subnetwork.

$$isa_{128} = \langle gna_{48} : sna_{16} : sap_{64} \rangle$$

An ISA has two accepted expressions, one a human-oriented text string referred to as a *universal resource locator* (URL) and one a binary (hexadecimal) string as defined above. URLs are typically in the form

service_protocol://service_host_name.domain_name.domain_extension

For example, a secure web (https) based CPS data acquisition (sensor) service might have a URL of the form

https://sensorID.cpsID.net

This URL would translate (through a Domain Name directory Service, DNS) into a specific Internet address of the form $isa_{128} = \langle gna_{48} : sna_{16} : sap_{64} \rangle$.

[5] Ref. http://www.ietf.org/rfc/rfc2460.txt

META COMMAND SYSTEMS ™
Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

Temporal References

Time, in infospace, is by definition relative to one or more network accessible (local or distributed) clocks. According to the prevailing Internet time standard[6], time in infospace consists of monotonically increasing offsets (timestamps) from January 1st, 1900, eight decades prior to the 1979 emergence of the internet *network time protocol* (NTP)[7]. Time according to NTP and its *simpler* non-averaging (SNTP) version[8], is concerned with the relation between times (durations) reported by different network-connected observers (CPS servers) in their description events of shared interest.

NTP time is synchronized with Coordinated Universal Time (UTC), the standard used nearly everywhere in the world, itself derived from a set of rationalized atomic clocks. Conceptually, UTC extends into the indefinite past and indefinite future. The NTP timescale is defined by a 128-bit register, of which the first 64 bits count seconds from the 0h January 1, 1900 (the *prime epoch*) and the last 64 bits count fractions of seconds. Timestamps are unsigned 64-bit fixed-point integers, with whole seconds to the left of the decimal point and left of the bit 32, numbered from the left (big-endian). Fractions of seconds are to the right of the decimal point. This format represents the 136 years from 1900 through 2036 with a precision of about 200 picoseconds.

As network bandwidth continues to increase and the processors in NTP client and server machines operate at super-gigahertz clock rates, issues of clock synchronization become critical for the proper coordination of high-resolution cyberphysical system applications. Such synchronization is increaslingly critic for the proper (e.g., *causal*) sequencing, recording and reporting of cyberphysical events (CPE).

Cyberspace-Time References

A cyberspatial object is one whose state requires description in both geospatial and infospatial terms. A dynamic cyberspatial object is one whose behaviors unfold in both cyberspace and time. With the preceding introduction to geospatial, infospatial and temporal consideration we now able to turn our attention to the behavior of cyberspatial objects and the set of services required of cyberspatial systems in their monitoring and control activities. Figure 11A describes graphically a region of cyberspace-time (CST) whose coordinates include a geospatial reference (GSA), and infospatial reference (ISA) and a timeline (NTS) established by internetwork timestamps. A specific CPS is assigned by an enterprise to this cell for the purpose of identifying, cataloging, monitoring and possibly controlling the dynamic properties of the cell.

The cyberspace-time axes are orthogonal since a value along any one dimension is independent of values along the others. Recall that the GSA and ISA dimensions are each three dimensional. CST is therefore a seven dimensional manifold with boundary constraints determined by the size and allocation of GSA and ISA address components.

Figure 11B shows the trajectory (motion) of a cyberspatial object (event, particle) as it moves in CST from point $P(t_0)$ to $P(t_3)$. The figure shows parametrically the movement of an object (event) through CST, where the sequence of network timestamps (along the NTS dimension) t_i is derived from NTP-synchronized clocks.

$$\{gsa(t_0), isa(t_0)\} \rightarrow \{gsa(t_1), isa(t_1)\} \rightarrow \{gsa(t_2), isa(t_2)\} \rightarrow \{gsa(t_3), isa(t_3)\}$$

We are interested in variations in this trajectory as perceived by different observers (CPS reference frames). These distinct reference frames arise from the relative cyberspatial locations of observers that are focused on specific regions of cyberspace-time.

---

[6] Ref. RFC-1305, http://www.ietf.org/rfc/rfc1305.txt?number=1305
[7] For a discussion of the history and current status of NTP, see http://www.ntp.org/
[8] Ref. RFC-4330, http://www.ietf.org/rfc/rfc1305.txt?number=4330

Meta Command Systems, Inc.
1045 W. Glen Oaks Lane, Suite 202
Mequon, WI 53092-3477 USA
+1.262.240.2956 (v); +1.262.240.2958 (f)
info@metacomsys.com; www.metacomsys.com

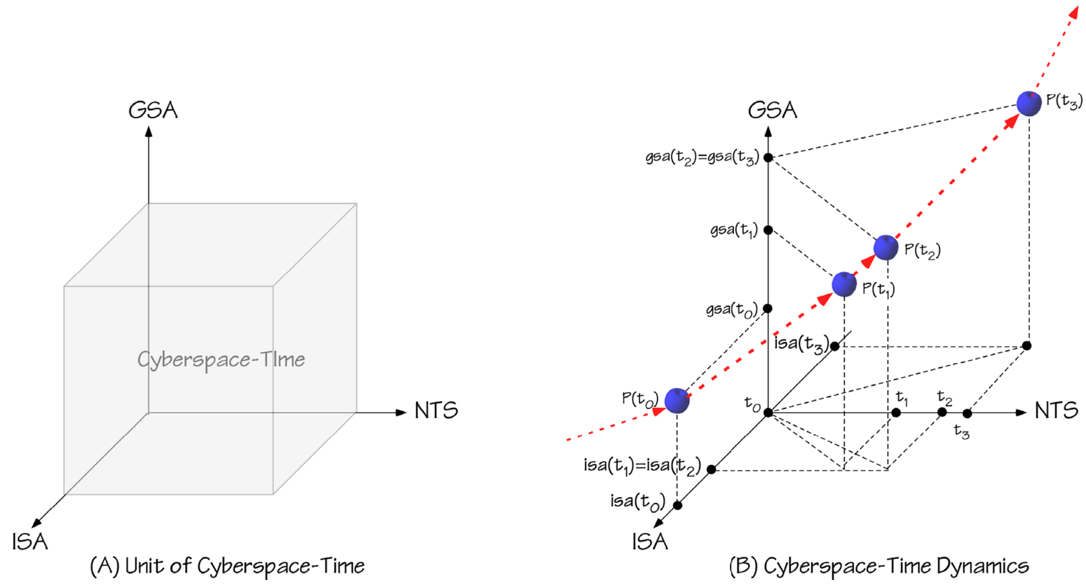


Figure 11 – Cyberspace-Time

In summary, Figure 12 shows our infospatial observer (agent) monitoring a region of cyberspace. Many details have been omitted for brevity.

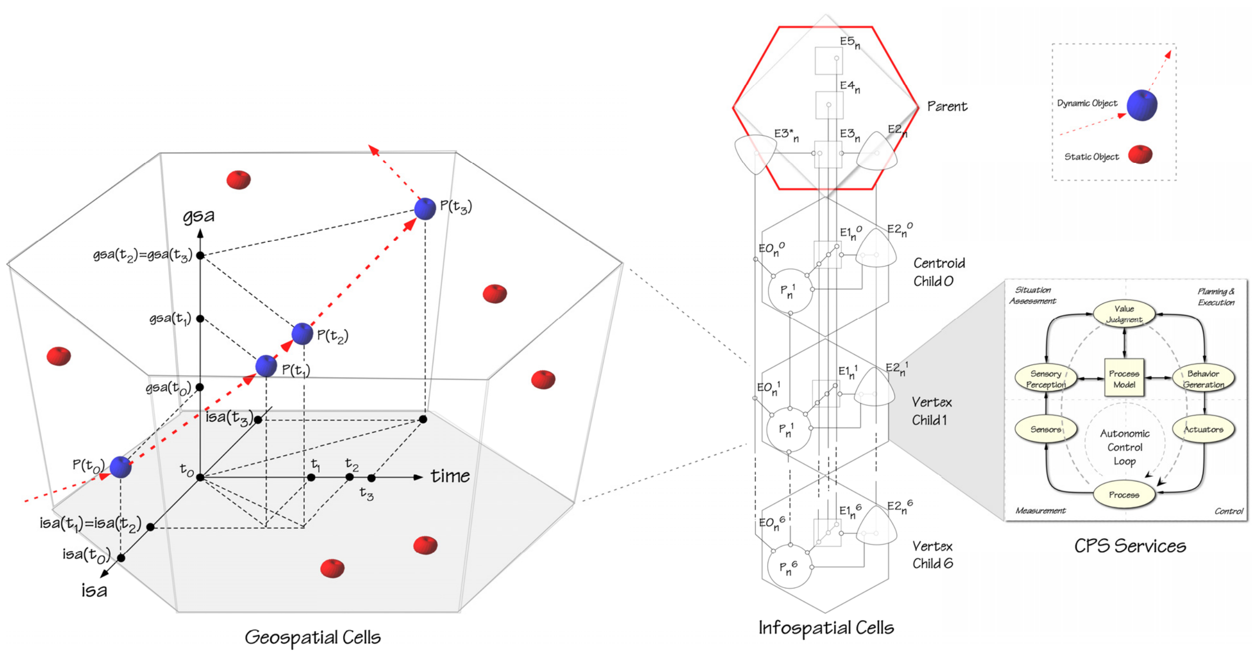


Figure 12 - Agent (Observer-Controller) Assigned to a Region of Cyberspace-Time

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Mark Birkin
School of Geography, University of Leeds

Multi-agent systems, mathematical modelling and microsimulation

I consider mathematical modelling and simulation to be both a legitimate intellectual exercise, and a practical aid to policy analysis and decision-making. However to the extent that there is an issue revolving around understanding models versus understanding systems[1], then I am fully committed to the systems camp. I think it is a cop out to produce models which may exhibit all kinds of interesting behaviour under idealised conditions which may bear little or no resemblance to real systems. It concerns me that there is a growing swell within the MAS community which perhaps regards questions like validation and policy application as faintly grubby and beneath the dignity of the simulation modelling community. These days it is easy to generate models which do all sorts of exciting things, and to visualise these models in novel and interesting ways; but it remains as difficult as ever to develop models which give real insights about real systems.

I am excited by the capabilities of e-Research to provide modellers with access to unprecedented riches of both data and simulation power. I think these opportunities are largely being ignored by the academic community, which remains too easily satisfied by proof of concept in relation to problems which are idealised, imaginary or trivial. For example, in relation to complex systems, there is too much rhetoric for my liking on the generation of complex behaviour from simple models as opposed to complex behaviour from complex models. It seems to me that much of the excitement about agents is in the ability to build models with very complex behaviour from agents with very simple behaviour. As geographers, I believe that the agents in our systems of interest actually have quite complex behaviours, whether those agents are consumers, regulators, entrepreneurs or whatever. I think we should be focusing more on realistic social simulations which recognise the existence of complexity throughout the (modelling) process.

One of my main methodological interests is in microsimulation. This technique is used in a number of large, policy-focused applications, many of them economically driven, although ‘spatial microsimulation’ has been emerging as a distinct research sub-theme. In these models, individuals and households are represented in great detail, in effect as a list of characteristics. These lists are typically generated either by reweighting survey sources, or by synthetic estimation from aggregate data. Behaviours can be added to synthetic microdata, for example by linking to meso-level spatial interaction models[2].

---

[1] Sayer R A, 1979, "Understanding urban models versus understanding cities" *Environment and Planning A 11(8),* 853 – 862.
[2] K. Hanaoka & G. Clarke (2007) Spatial microsimulation modelling for retail market analysis at the small area level, *Computers Environment and Urban Systems, xxx* (in press).

Models with this structure can begin to address difficult aggregation problems, but there remains a suspicion that they are overly deterministic. I am intrigued by the possibility that the incorporation of agent driven behaviours within a microsimulation model could somehow dramatically enrich the representation of dynamics at the level of individuals and households. I think the workshop would be a good opportunity to explore whether there are fundamental methodological mis-matches between the microsimulation and agent-based approaches to modelling, or whether these differences are more cultural and terminological.

I guess there is another question here about data, which is whether the two modelling styles – microsimulation and agents – actually require different types of data, with the former focused on 'statistical data' of the type readily available from censuses and major surveys, the latter needing data of a more 'behavioural' orientation, and consequently more difficult to access. This looks to present overlaps into disciplinary and methodological domains like sociology, anthropology and ethnography, although experience indicates that geographers are well-placed to exploit such multi-disciplinary terrain.

These are very real practical concerns, as it is our intention to combine microsimulation with agents in the Moses project. For example, consider the problem of the impact of an ageing population on the provision of social services. From a microsimulation perspective, we can look at this problem in terms of transition probabilities, from young to old, from healthy to sick, from married to single or widowed, and so on. But what effect does something like a social network have on this process, given that the vast majority of social care is provided informally within the context of families and neighbourhoods? Is this something that should be represented with consideration of the interaction between individuals as agents?

Another important concern is with the robustness of forecasts from simulations, whether agent-based or otherwise. Since I have argued that policy-relevant models need to be calibrated to extensive data sets, this presents obvious problems to predictive modelling where 'data' relating to the future is clearly much more difficult to come by. Although there are many potential strategies for the evaluation of forecasting capabilities, such as historical benchmarking (calibration of the model to a historical baseline, so that model 'forecasts' can be evaluated against subsequent events that are known), model training strategies (in which a portion of data is withheld for model evaluation), continuous monitoring, or even running the models in reverse, none of these mechanisms appears to be completely satisfactory.

I have been impressed by a number of high profile simulations of the spread of diseases through spatial networks, notably those emanating from UCL[3] and from Los Alamos,

[3] Neil M. Ferguson, Derek A.T. Cummings, Simon Cauchemez, Christophe Fraser, Steven Riley, Aronrag Meeyai1, Sopon Iamsirithaworn & Donald S. Burke (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia, Nature, 437 (8), 209-214.

now VBI[4][5]. These examples help to establish the credibility of agent models, and the emphasis on real policy applications is particularly welcome. It seems to me that the application of these methods has a much broader relevance to problems of diffusion in space in time [6] – indeed there is direct resonance with our own humble and much less sexy analysis of spatial patterns and retail price dynamics[7]. However the naivete of the spatial networks which underpins these models is alarming. I feel sure there is a whole apparatus of symmetrical nodes, radial accessibilities, well-regulated hierarchies, and all sorts of traditional and unsatisfactory representations beneath these models. It seems to me to be important that as geographers and spatial scientists we can get a message across to a broader MAS community that geography matters, and the workshop would be a useful opportunity to discuss ways to promote this.

---

[4] Stephen Eubank[1], Hasan Guclu[2], V. S. Anil Kumar[1], Madhav V. Marathe[1], Aravind Srinivasan[3], Zoltán Toroczkai[4] and Nan Wang[5] (2004), Modelling disease outbreaks in realistic urban social networks, *Nature, 429,* 180-184.

[5] Chris L. Barrett, Stephen G. Eubank and James P. Smith (2005) If Smallpox Strikes Portland ... , *Scientific American, 292 (3),* 54-61.

[6] Bo Lenntorp, Gunnar Törnqvist, Olof Wärneryd, Sture Öberg (2004) Torsten Hägerstrand 1916-2004, *Geografiska Annaler, Series B: Human Geography 86 (4),* 325–326

[7] Heppenstall, A.J., Evans, A.J., and Birkin, M.H. (2006) Application of multi-agent systems to modelling a dynamic, locally interacting retail market, *Journal of Artificial Societies and Social Simulation, 9, 3*.

Position Paper for Agent-Based Models of Complex Systems*

Dan Brown, Professor
School of Natural Resources and Environment
University of Michigan

Object-oriented process models, which include individual-based models (IBM) commonly used in ecology and agent-based models (ABM) common in the social sciences, allow for modeling both change and movement of geographic entities and have developed independently of GIS. The object-oriented framework of ABM involves identification of agents and of a temporal framework within which those agents perform actions. While many different types of agents can exist, the following general definition is common: an agent is a self-directed object, i.e., it has the ability to satisfy internal goals or objectives through actions and decisions based on a set of internal rules or strategies. These agents may be dynamic in either state (i.e., change) or space (i.e., movement) and may, through their actions, change the state or location of other objects, processes, or environments around them. Agent dynamics are most naturally implemented in an ABM by a set of behaviors ("methods") that can include conditional decision making and other (non-linear) rules that distinguish them from mathematically continuous models. The ability for Lagrangian motion (i.e., agent movement) distinguishes ABM and other object-oriented modeling frameworks from the change-based spatial models described above. It also creates additional challenges for integrating these models with GIS, as described in more detail below.

ABM dynamics are defined at the level of (a) agent behaviors that result in change and movement, and (b) the independent dynamics, if any, of non-agent objects. Thus to represent dynamics, ABMs are implemented as discrete event simulations, in which some kind of "scheduling" mechanism handles the sequencing of agent behaviors and events. An ABM may implement scheduled events in three ways:

- Events may be sequenced in a synchronous step-wise fashion. For example, each agent, set of agents or non-agent object is signaled to perform its tasks once at each time step or once every *n* time steps.
- An event may be scheduled to occur only once at some time step *n*. Any number of different events may be scheduled to occur in this fashion providing a predetermined history of events to take place.
- The model may encapsulate 'event-driven' processes whereby model agents may trigger events to occur or may add events to the schedule or queue of events to take place.

On the other hand, ABMs often use relatively limited representations of space. For example, ABMs frequently use hypothetical spaces based on square or hexagonal tessellations, and only recently have ABMs begun to use real-world spatial data. To avoid edge effects on the performance of some models, researchers commonly use a toroidal representation of space, which wraps around from top-bottom, left-right, and vice versa. The rich temporal representations (agents and processes) of agent-based models, therefore complement the spatial data representations (fields, objects and functions) of GIS. The object-oriented nature of both presents tremendous opportunities for their integration.

Given the complementarities of spatial data models (fields and objects) and agent-based (i.e., object-oriented) process models, and their combined potential to improve on integrated representations of spatial patterns and temporal processes, we argue that tight coupling of models and data within ABM and GIS, respectively, can reap benefits in terms of both efficiency, through reduced computing times, and capability, through new functionality. Attempts to integrate ABM and GIS techniques have raised several conceptual and technical questions. These issues broadly fall into questions of ontology and process, i.e., how are entities and processes represented, and how do those representations interact, respectively. For instance, Bian (2003) concluded that the environment within an individual-based model can be represented as either patch-based (i.e., object-based), maintaining object-orientation in both the model and data, or field-based, such that object-oriented individuals interact with a discretized environment of attributes. She discounts the value of treating all cells in a grid-based environment as objects on both technical (i.e., due to inefficiencies) and ontological (i.e., poor match to conceptual view of fields) grounds.

More generally, developing models that make use of both GIS and ABM techniques requires the specification and implementation of relationships between agent-level processes and spatial data. First, by defining an *identity relationship* between an agent and a spatial feature or features, GIS techniques can be used to store the geographic extent and attributes of the feature, while ABM techniques represent the behavior of the agent and the change in associated feature(s). Thus (a) spatial features associated with agents can move or change, and (b) attributes of features associated with agents can change. Second, in many models, agents have *causal relationships* with (i.e., the ability to take actions that affect) spatial features and/or their attributes, even if there is no identity association between the agent and the spatial feature(s) it is acting on (i.e., non-agent features). Agents can take actions that result in changed locations or attributes of features, or they can take actions that change the values of an attribute on a field (e.g., a raster). Third, *temporal relationships* are inherent in two types of actions in a coupled process-data model: (a) the actions of the agents and (b) the updating of attributes or locations of features in a database or display. Either can be handled using synchronous or asynchronous approaches. Finally, movement of spatial features, either by processes internal to their associated agents or by those of other agents, can require basic information about the *topological relationships* between an agent and the physical world or between features.

We have been working on a number of different approaches to implementing the object-based process models that are linked to dynamic spatial data models. The simplest approach has been to loosely couple agent-based process models, with GIS data bases by passing interchange files between the two. A significant disadvantage of this is the volume of data created by the models and the consequent data-management challenges. Secondly, we participated in the testing of AgentAnalyst, an extension to both RePast and ArcGIS and other GIS systems (like OpenGIS) that serves as a sort of middleware to link the two. There are limits to its ability to dynamically use of GIS functions, unless connected to open source GISs. This places greater burden on the modeler to program GIS functions within the model. Finally, we have a model written in VBA and running completely within ArcGIS, which takes fuller advantage of GIS functions, but requires that ABM functions (like the scheduler) be programmed into the model (rather than relying on existing software libraries). New dynamic modeling functions within GIS environments, like those in PCRaster for raster data, will significantly improve these capabilities within next-generation GISs.

* mostly excerpted from Brown, D.G., Riolo, R.L., Robinson, D., North, M., and Rand, W. Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems,* 7(1): 1-23.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Crispin Cooper, CPLAN, University of Cardiff

Statement of interest

February 15, 2007

Evolution and Development

My professional background is in genetic algorithms, which I have in the past applied to finite element models and electronic circuits e.g. (Cooper et al. 2006). However, in future I aim to apply similar techniques not to design new technologies, but to assist in the understanding of the social, spatial and economic world around us.

In particular, I have an interest in developmental computation - the process of growth which translates the genotype (the component of the system which evolves), into the phenotype (the component which actually survives or perishes). I believe this process to be essential to the evolution of non-trivial structures. As the *size* of any evolved structure increases, the *search space* of possible structures grows exponentially, so the probability of a given phenotype emerging by genetic processes alone rapidly approaches zero. It is not useful to examine natural and social systems and merely state that they are 'very unlikely'. Growth (as nature's answer to this so-called complexity crisis) is a crucial component in the creation of such systems, and therefore crucial to our understanding of them.

It should be noted that the processes of growth, and evolution, have themselves evolved. To talk usefully about such concepts in the real world we therefore need to step outside the narrow ontology of genetic algorithms (which view most systems as a process of mutation, evaluation and selection) and instead employ ideas from the world of artificial life, which views evolutionary processes as complex emergent systems. I consider Ray (1991) to be a key work in this area; it is a generative study in which computer programs compete with one another for resources and an artificial ecology emerges without any external fitness function. However, traditional complexity science probably has a part to play in formalising the ideas behind this, and game theory is also useful when/if we can assume that agents behave rationally.

Networks and other formalisms

Network formalisms (e.g. random, small-world and scale-free networks) have much potential for modelling the rich levels of interaction present in the real world, as they allow us to study chains of cause and effect which may not relate to spatial proximity - this is of course one of the advantages which agent based models offer us over and above cellular automata. Network formalisms are still a young field, as is shown by the comparatively recent work on for example, weighted networks (Latora & Marchiori 2003) and correlations between network topology and real-world parameters (de Montis et al. 2005). I view this type of study as essential if network formalisms are to be applied to understanding the world around us.

For me, this field also serves to highlight another question: how else can we usefully formalize real-world structures with mathematical abstractions? Network theory seems to be only one of many possibilities; for example, Markov chains have also recently been applied to studying evolutionary processes (Wheeler 2006).

Applications: real tools from a virtual world

Originally my doctoral research proposal (for the Cardiff University School of City and Regional Planning) concerned the agent-based modelling of urban growth; however the scope has widened somewhat. I am currently seeking problems in the fields of spatial economics, economic geography and evolutionary economics. An example is Tassier & Menczer (2001) in which a realistic social network (of employment contacts) emerges by evolutionary processes alone, and is then analysed from a perspective of efficiency. Vromen (2006), as another example, takes an interesting perspective on growth as it relates to the expansion of knowledge, by suggesting that the evolving 'genotype' consists of skills and behavioural routines, while the 'phenotype' consists of the ideas formed by these entities interacting with their environment. However, this is a discursive argument rather than the result of experimental simulation.

Both of the above studies do not deal with the spatial dimension; however, I think plenty of potential exists not only in analysing the *virtual* space of network formalisms, but also in analysing its relation to *real* physical space.

I see two main uses for computer simulation: it can either be seen as a calibrated modelling system, which aims to predict future real-world developments with a high degree of accuracy; or as a virtual world which isn't expected to match reality precisely but which nevertheless can be used to gain insight into generic processes. Except in a few special cases such as traffic simulation, I am skeptical as to the power of the former approach; therefore I am more interested in the study of generic processes, classed as 'theoretical' in Wu (2005). I hope that this will in turn lead to the development of tools which can be applied to the analysis of real data.

References

Cooper, C., Murphy, D., Howard, D. & Tyrrell, A. (2006), 'Singing synthesis with an evolved physical model', *IEEE Transactions on Audio, Speech and Language Processing* **14**(4).

de Montis, A., Barthelemy, M., Chessa, A. & Vespignani, A. (2005), 'The structure of inter-urban traffic: A weighted network analysis'.
URL: *http://www.citebase.org/abstract?id=oai:arXiv.org:physics/0507106*

Latora, V. & Marchiori, M. (2003), 'Economic small-world behaviour in weighted networks', *The European Physical Journal B* **32**, 249–263.

Ray, T. S. (1991), *Scientific Excellence in Supercomputing: The IBM 1990 Contest Prize Papers, Athens, GA*, The Baldwin Press, The University of Georgia, chapter Evolution and optimization of digital organisms, pp. 489–531.

Tassier, T. & Menczer, F. (2001), 'Emerging small-world referral networks in evolutionary labour markets', *IEEE Transactions on Evolutionary Computation* **5**(5), 482.

Vromen, J. J. (2006), 'Routines, genes and program-based behaviour', *Journal of Evolutionary Economics* **16**, 543–560.

Wheeler, S. (2006), 'A Markov analysis of social learning and adaptation', *Journal of Evolutionary Economics* **16**, 299–319.

Wu, F. (2005), *GeoDynamics*, CRC Press, Boca Raton, Florida, pp. 205–214.

Cellular automata and agent-based models: what next?

Helen Couclelis

Cellular automata (CA) and agent-based models (ABM) are hallmarks of computational geography. Increasingly they are used in combination in the development of process-oriented models of people or other kinds of actors interacting with their environment, with CA typically simulating the spatial environment and ABM representing the relevant decision-making units. Combined CA-ABM models may be used to simulate farmer communities dynamically affecting land cover through adaptive land use decisions, households seeking suitable housing in changing urban areas, or trip makers responding to various congestion pricing schemes on transportation networks. Such models are built to incorporate considerable empirical and intuitive understanding of the complex processes of interest, and when calibrated to actual data they are often presented as suitable for prediction and policy analysis. As someone who has studied these two types of models for over 20 years, I am skeptical of such claims (Couclelis 2001). I believe that we have by now accumulated enough experience for a more systematic exploration of the potential and the limitations of CA and ABM to model complex spatial systems, whether used in conjunction or separately.

Computation has often been discussed as the third way of doing science, lying somewhere between theory development and experimentation. This implies a new approach to knowledge production and the need for a new kind of research methodology different from either the mostly deductive mode of theoretical work or the mostly inductive mode of experimental science. That third way centers on the construction of complex simulated worlds within which experiments may be run that would have been difficult or impossible to conduct in the real world. The epistemological problem is that models of complex open systems with deep uncertainties, as social systems nearly always (and natural systems usually) are, cannot in principle be used for prediction. Predictive models belong in the traditional scientific paradigm of theoretical closed system descriptions supported by experimental evidence, or at least of well-established empirical generalizations such as human geography’s spatial interaction models. Because this fact is not always appreciated, many computational modelers understand progress in the field to mean building simulations that are increasingly detailed and realistic, even though increased detail can actually decrease any predictive value such models may have. The meaning of model validation in this new world of computational process models thus remains open, and so does the question of how to derive valid insights that may be useful for both theory development and for policy guidance.

Technically, the reason why models of complex open systems cannot yield reliable predictions is that many (in some cases infinitely many) different models can provide acceptable fits to the data. In other words, any particular model is but one realization out of a large space of potential models, few or none of which may be correct by whatever definition of the term. This issue is sometimes addressed with Monte Carlo simulations that generate many versions of a particular model by systematically varying the parameters; model outcomes are then considered reliable to the extent that they are reproduced by large numbers of different parameter sets. This methodology may take care of parametric uncertainty but cannot address structural uncertainty, that is, the degree of confidence one may have in the structural validity of the model. Researchers in both the social and the natural sciences have suggested methods for generating large numbers of different model structures in a manner analogous to generating versions of the same model through Monte Carlo simulations. The idea is that investigating the properties of entire ensembles of models, even relatively simple ones, may yield more robust insights into the complex spatial processes of interest than the study of even the more realistic-looking individual models. Procedures for generating ensembles of models for that purpose have been described in hydrology by Beven and associates in a long series of papers (e.g, Beven and Freer 2001), in policy studies by the RAND team of Popper, Lempert and Bankes (e.g., Popper et al. 2003), and in several other fields.

Should we wish to explore that direction, our task will be greatly facilitated by the fact that in formal terms, CA and ABM are very close cousins. Both are structures described in the theory of automata, one of the three major branches of the mathematical theory of computation. A CA may be seen as spatial array of ABM. In principle, anything that can be modeled as a (generalized) CA can also be modeled as an ABM and vice versa, though obviously some options will be more intuitive and computationally efficient than others. Thus CA models have been developed where the cells are endowed with complex goal-directed decision rules and ABM where the agent is the environment. Some researchers consider mobility to be the defining difference between the two kinds of models, but in actual fact CA simulate movement in the same way your computer screen does, by spreading activation down a sequence of adjacent cells or pixels. (Action at a distance – easy for ABM – is somewhat trickier to simulate within a pure CA framework, but that too can be done). The affinity between ABM and CA means that both agents and environment can be specified within the same framework in the formal language of automata theory. Such integration is very likely to provide a substantially increased theoretical understanding of the properties of these structures and to greatly support the generation and analysis of appropriate ensembles of models. I think that there is fertile ground here for the more theoretically inclined among us to make contributions to complex spatial systems modeling that could benefit researchers in many fields.

References

Couclelis, H. (2001). Why I no longer work with agents: a challenge for ABMs of human-environment interaction. In D. Parker, T. Berger, and St. Manson (Eds.) *Agent-based models of land-use and land-cover change*, LUCC Report Series No.6, 3-5.

Beven, K. and J. Freer (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249 (2001), 11-29.

Lempert, R.J., S. Popper and S. Bankes (2003). *Shaping the next one hundred years: new methods for quantitative long-term policy analysis.* RAND, Santa Monica, CA.

Representing geographic dynamics

Thomas J. Cova, Associate Professor, Department of Geography
University of Utah, 260 S. Central Campus Dr., Rm. 270,
Salt Lake City, UT 84112-9155, cova@geog.utah.edu, Voice: 801-581-7930

A number of related yet distinct subfields of GIScience associated with representing geographic dynamics have recently emerged. Researchers in spatio-temporal theory and data modeling, spatial process modeling, complex systems, and agent-based modeling all consider dynamic (or temporal) aspects of geographic phenomena germane to their respective specialty. Given this common ground, an obvious question arises as to the degree to which these communities have shared goals that might be forwarded through collaboration. In short, where might synergies or unique insights lie, and what redundant efforts might be streamlined? This short position paper looks at two research areas (and communities) that are the focus of this workshop: agent-based modeling and spatio-temporal data modeling.

Agent-based modeling of complex systems

Agent-based modeling is an active research area that continues to gain momentum. The software tools to support this type of modeling have come a long way both in terms of facilitating rapid model development as well as integrating models with GIS. The area also evolved from the start with a particular emphasis on explicit representation of space despite having no relation to the GIScience community. In general, researchers in this area are concerned with modeling geographic processes "from the bottom up" in a quest to understand emergent phenomena at a macro scale form micro interactions. A common approach is to generate *synthetic* populations and landscapes where autonomous, intelligent agents can interact with their environment and each other. Although the landscapes represented are often based on real places, the notion that the population is synthetic implies that the correspondence with actual agents in the real world is secondary to the goal of discovering emergent outcomes at the macro scale. For example, it's more important to reveal global patterns of segregation in a hypothetical city than it is to represent any real agent, building or event. In some cases, agents may adapt to, or learn from, their surroundings. In this way, agents are typically ephemeral. Validity and accuracy therefore generally refer to the degree to which a model corresponds with a given process rather than whether any object, interaction or event is positioned correctly in space-time (or can be recovered through query).

Spatio-temporal data modeling

In contrast to agent-based modeling, research in spatio-temporal data modeling tends to focus on accurately depicting real world geographic phenomena as objects, fields, events and processes in such a manner that their histories and interaction can be queried, reconstructed or predicted. Mining large data sets that represent spatio-temporal phenomena is a current theme. Maintaining or tracking an object's identity through time is also important, as is assessing the accuracy of a field or object attribute at time t. An example object in this research context might be a delivery truck traveling across the country, a field might be the anticipated precipitation along the route, and an event might be "leaving Iowa" or "entering a storm cell". Field-object hybrids have been proposed to deal with unique phenomena like weather or wildfire (Yuan, 2001; Cova and Goodchild 2002) along with models that focus on the inter-relationships between objects, events, and processes (Worboys and Hornsby 2004). Recent work has emphasized the search for an underlying theoretical level that might underpin all geographic representation (Goodchild et al. 2007). In any case, the concepts of synthetic populations, emergence, complexity and adaptiveness are generally foreign to this area, despite the similar focus on geo-dynamics shared with researchers in complex systems and agent-based modeling.

Combining efforts

What would a union between the sub-fields of complex adaptive systems (or agent-based modeling) and spatio-temporal data modeling yield? One outcome might be a subfield (or associated software platform) that can model adaptive agents and their interactions with an eye to identifying emergent phenomena, yet also allow for greater data mining or querying of object and field dynamics or events to reconstruct scenes with a close (or accurate) relationship with real world states. The theories and software tools in the two areas seem to have developed from entirely different goals, and it's difficult to visualize one platform that would satisfy specialists from both areas. A recent call from the point of view of spatio-temporal theory for research into data structures (Galton 2004) emphasizes how far behind this area is from the easy-to-acquire and install software platforms to support agent-based modeling. Similarly, recent papers from spatial-temporal data modeling tend to emphasize how difficult it is to advance this area given the problems that time presents (Peuquet 2001; O'Sullivan, 2005). The notion of a workshop to explore the benefits of cross-fertilization of these two areas is likely to yield very interesting and fruitful research and development directions for both areas.

References

Cova, T.J. and Goodchild, M.F. (2002) Extending geographical representation to include fields of spatial objects. International Journal of Geographical Information Science, 16(6), 509-532.

Galton, A. (2004). Fields and Objects in Space, Time, and Space-time. *Spatial Cognition and Computation*, 4(1), 39-68.

Goodchild, M.F., Yuan, M., and Cova, T.J. (2007) Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3), 239-260.

O'Sullivan, D. (2005). Geographical information science: time changes everything. *Progress in Human Geography*, 29(6), 749-756.

Peuquet, D. (2001). Making Space for Time: Issues in Space-Time Data Representation. *Geoinformatica*, 5(1), 11-32.

Worboys, M.F. and Hornsby, K. (2004). From Objects to Events: GEM, the Geospatial Event Model. *Lectures Notes in Computer Science*, 3234, 327-343.

Yuan, M. (2001). Representing complex geographic phenomena in GIS. *Cartography and Geographic Information Science*, 28(2), 83-96.

Workshop on Agent-Based Modeling of Complex Spatial Systems
14-16th April 2007

Catherine Dibble
Assistant Professor, Department of Geography, University of Maryland cdibble@umd.edu

Agent-Based GeoComputation is not about building models. It is, ultimately, about the art and science of addressing complex spatial problems with computers; for sound inference on matters of importance. An appropriate model is necessary, but not sufficient. Consider the complementary laboratory tools and practices sufficient for inference with agent-based computational laboratory research for effective responses to complex geographic emergencies.

Our methodological advances in computational laboratory tools and practices have been driven by a potential pandemic influenza emergency due to an unusually deadly (H5N1) influenza pandemic spreading worldwide among migrating, wild, and domestic birds. H5N1 influenza currently kills more than half of its human victims. If it adapts to spread easily among humans, it could lead to a global influenza pandemic far more severe than the flu pandemic of 1918-1919.

Our research uses complementary computational laboratory tools to evaluate relative geographical risks via patterns of inter-city hierarchical diffusion of pandemic influenza, and to optimize geographic deployment of limited resources to inhibit the inter-city spread of pandemic influenza. The crucial substantive challenge harnesses advantages of geographic structure to clarify relative risks and to amplify the protective leverage of available resources, in order to target advanced preparation, minimize mortality, and gain time for vaccine production and administration. The crucial methodological challenge is to conduct sufficient sensitivity and risk analyses of the model and of recommended interventions to provide as much information as possible for decision-makers, in the event that model results must be used to recommend deployments of interventions during a pandemic emergency.

The epistemological insights and methodological extensions driven by responding to this potential emergency provide helpful guidelines for effective calibrations, experimental designs, optimizations, and risk analyses of spatial agent-based computational laboratory research.

Complementary Mathematical Modeling of Complex Dynamic Systems

Although respective limitations and advantages of mathematical versus agent-based models of complex dynamic systems are most commonly considered with respect to their roles as methodological substitutes (see for example Rahmandad and Sterman (2004)), there exist significantly complementary roles for which understanding their differences is also vital.

First, hybrid models using Dynamic Agent Compression (Wendel and Dibble (2007)) can harness efficiency and scaling gains by using mathematical equations rather than representative agents. For which temporal nuances of sacrificed discretization may become important to understand even for generalizations that appear to be lossless in other ways.

Second, standard mathematical models of complex dynamic systems such as SEIR models of epidemic diffusion in aspatial populations (Anderson and May 1991) may be cast as official standards against which behavior of agent models should be benchmarked. To do so, however, requires explicit consideration of the agent-based effects due to discretization as well as of those due to spatial structure in agent models (Durrett and Levin (1994), Keeling and Grenfell (2000), and Rahmandad and Sterman (2004)).

Complementary Genetic Algorithm Inference, Optimization, and Risk Analysis

Miller (1998) proposed to use a supervisory genetic algorithm to perform what he called "active nonlinear tests" (ANTs) by using the genetic algorithm to challenge each simulation model by seeking outcomes that provide exceptions or counter-examples to its usual results. This section briefly discusses a generalization of Miller's ANTs to the broader problem of providing effective search and optimization across both treatment domains and outcome ranges for a model.

Systematic analysis of model behavior typically involves millions of simulation runs, each controlled by sweeping across discrete lists of values for sensitive model parameters, for each treatment variable of interest, and for seeds to control one or more random number series for stochastic simulations.

Of far greater importance scientifically, the standard focus on exploring model behavior via combinatorial sweeping across regularly spaced parameter values is a blind search for significant outcomes. As illustrated in Figure 1, regularly spaced parameter values may be entirely unrelated to the truly important parameter values where model outcome reach significant extrema.

Ideally, we would like to be able to search for interesting behavior in the outcome space rather than sweeping blindly in parameter space. As illustrated in Figure 2, a supervisory genetic algorithm allows us to do so with far greater efficiency than brute force combinatorial sweeping of parameter spaces. The genetic algorithm can be set up to search across combinations of key parameters for extreme values of single or multiple combinations of outcome variables, based on results from one or more stochastic replications of the scenario that is associated with each combination of key parameters. In addition, the greater economy in searching for key scenarios releases computational resources that may in turn be used to simulate sufficient stochastic replications for each to be able to distinguish statistically significant differences among scenario outcomes.

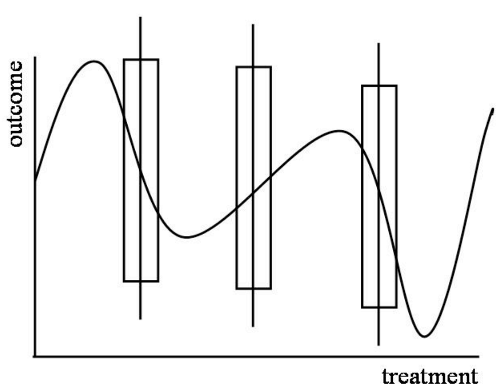


Figure 1: Running only a few stochastic replicates of each treatment level can result in variances so large that the signal becomes lost in the noise. Similarly, selecting treatment levels blindly via random or regular spacing may completely miss important local and global extrema.

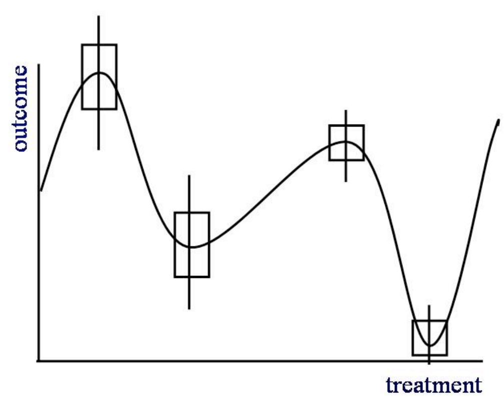


Figure 2: In contrast, an ideal experimental design runs enough stochastic replicates for reliable inference. Similarly, data-driven experimental designs may provide guidance for identification of key values for treatment variables and for basins of attraction leading to common outcomes.

References

Anderson, R. M. and R. M. May (1991) *Infectious diseases of humans: Dynamics and control*, Oxford: Oxford University Press.

Dibble, Catherine (2006) Computational Laboratories for Spatial Agent-Based Models, Ch 31 in: Tesfatsion, Leigh and Kenneth L. Judd, Editors, *Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics*, Amsterdam: Elsevier.

Dibble, Catherine and Philip G. Feldman (2004) The GeoGraph 3D Computational Laboratory, *Journal of Artificial Societies and Social Simulation* vol. 7, no. 1, http://jasss.soc.surrey.ac.uk/7/1/7.html

Durrett, R. and Levin, S. (1994) The Importance of Being Discrete (and Spatial), *Theoretical Population Biology* 46(3):363-394.

Keeling, Matt J. and Bryan T. Grenfell (2000) Individual-based Perspectives on R_0, *Journal of Theoretical Biology* 203:51-61.

Miller, John H. (1998) Active Nonlinear Tests (ANTs) of Complex Simulation Models, *Management Science* 44, no. 6: 820-830.

Rahmandad, H. and J. D. Sterman (2004) Heterogeneity and network structure in the dynamics of contagion: Comparing agent-based and differential equation models, Submitted to *Management Science*.

Stanford's *Biomedical Computation Review* (Summer 2006). (The article is about MIDAS projects; our work is mentioned in its final two pages.) http://biomedicalcomputationreview.org/2/3/7.pdf

Wendel, Stephen and Catherine Dibble (2007) Dynamic Agent Compression, *Journal of Artificial Societies and Social Simulation* vol. 10, no. 2, 9, http://jasss.soc.surrey.ac.uk/10/2/9.html

Bridging the Gulf between ABM and CSS: A Three-Tiered Approach

A Perspective for the Workshop on Agent-Based Modeling of Complex Spatial Systems

Hamid R. Ekbia
hekbia@indiana.edu

That there is a gulf between the agent-based modeling (ABM) and complex spatial systems (CSS) communities is hardly in need of elucidation. Part of the reason for the gap is historical and institutional — these areas originate in different disciplines and are practiced in different quarters of the academia. Add to this the fact that "space" has only recently started to acquire its due status, along with "time," in our thinking, and one could readily explain the gulf. To connect the two communities, therefore, needs particular institutional efforts and arrangements, of which this workshop is an example. In addition, however, there are other gaps between the two communities, which I would like to characterize as conceptual, methodological, and technical. To bridge the gulf, I suggest, work needs to be done on all three levels. The following is a sketch of my thoughts on each.[1]

The Conceptual Level

The conceptual gap has multiple dimensions, but I would like to focus on how the two communities understand the core concepts of "complexity" and "representation."

Complexity

Complexity is an overused (and these days even abused) term. It means different things to people from different backgrounds and disciplines. For some (e.g., mathematicians and computer scientists) it has to do with quantity, scale, and magnitude, for others (e.g. psychologists and cognitive scientists) with structure, heterogeneity, and interconnection, and for yet others (e.g., biologists) with history, change, and function. The ABM and CSS communities, due to their multidisciplinary make-up, might not fully align with any one of these camps, but I believe that they do have varying understandings of "complexity." To illustrate this let me use a famous example.

Many decades ago, Herbert Simon made a simple observation, which is probably one of the most frequently cited episodes in modern science. The casual observation had to do with an ant's movement on beach sand, which Simon used to show how apparently complex behavior would emerge from the interaction of a simple organism (the ant) and a complex environment (the patterns of sand). Confronted with this scenario, the ABM community would mainly see the ant, and the CSS community would probably focus on the sand. I might be oversimplifying here, but the example highlights the differences in perspective.

Representations

Representations are also understood differently by various disciplinary traditions — e.g., as surrogates, precursors, and pointers to action, as mediators for coordination among different actors, as channels of communication, as vehicles of conflict resolution and alliance formation, and so on. Traditional accounts of representation typically focus only on the role of representations as stand-ins for individual activity, and ignore other equally important roles that they play in collective processes. Furthermore, they are based on an epistemological view that takes representations as products of a mapping between an external reality and an internal

---

[1] These thoughts mostly derive from ongoing research on the modeling of change and movement in GIS that I have been conducting in collaboration with the Redlands Institute.

mechanism. This limits the applicability, if any, of such accounts to the very special case of local, stable and static situations, where individual activity is the focus of attention, and where explicit linguistic forms of representation are dominant. What we are increasingly observing, on the other hand, is a whole set of collective practices, mostly mediated by modern digital information and communication technologies, which involve tacit, distributed, and indirect forms of representation using various mediums of expression (e.g., visual). In short, there is a huge gap between current narrow accounts and the broad aspects of representation.

Here again, the ABM and CSS communities might be wedded to one or the other of the above views of representation, and they need to develop a shared understanding of representations by paying attention to the increasingly multifaceted role that they play in the coordination of activities among temporally, geographically, and semantically dispersed actors.

The Methodological Level

What I call the methodological gap has mostly to do with the way the two communities approach and implement the phenomenon of "change." Traditionally, GIS views the world as a collection of locations and/or objects with attributes, and cartography views change as the application of rules to layers. Accordingly, the CSS community has a largely "snapshot" view of change as the implied difference between states ($S_1 - S_2 \Rightarrow \Delta s$), as opposed to the "incremental" (or "transitional") view that sees change as the accrual of effects in transition from one state to the next ($S_1 + \Delta s \Rightarrow S_2$)).

As others have shown, this difference in thinking about change has significant consequences, and the two communities need to make their differences as explicit as possible in order to be able to tackle common topics and issues.

The Technical Level

Finally, there are serious differences in terms of the computational techniques and programming environments applied in ABM and CSS. Traditionally, in GIS spatiotemporal information was represented by time-stamping records (data objects), attributes (fields), or attribute values (cells). Later on, there was a shift toward the integration of time and space via events or processes. Although semantically rich and more easily amenable to object-oriented modeling techniques, this integrated approach has proven to be non-trivial and challenging in many ways — e.g., in dealing with multiple scales (resolutions), in maintaining database consistency, and so on. More recently, there is a growing interest in agent-based modeling techniques, although the GIS community is yet to come to grips with agent-based modeling and to fully appreciate its potentials.

****

To sum up, I believe that there is great potential to be realized in interactions and collaborations between the ABM and CSS communities. In my own work on modeling change and movement in GIS, I have seen a lot of room for the integration of ABM techniques. But there are also serious barriers and challenges that need to be addressed on different levels.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Simulation, Noir

Andy Evans, Leeds

It was three ten in the morning when the smell of cordite and the memory of an unpleasant burlesque woke Jake Kadigin from an overwise delightful sleep. He'd been dreaming of dancing with a group of increasingly beautiful marionettes. Each dance had been of added intricacy, building to such a complexity that in the final moments before it all went sour there was little he could do but recline backwards into the cat's cradle of strings and be bound to the dancers himself. As his mobile phone implant injected unpleasant memories directly into his consciousness, the dancers became grotesques and the strings lit fuses, and what had started as a warm delight became a dark struggle from which he woke with a start of recognition. Must change that ringtone.

The call was from The Simulation. The Simulation had had other names over the years, but after the first ten pilot projects and the following thirty development studies, the group had run out of smartarse acronyms, and so the model, which spanned some six continents, had been worn down to just "The Simulation", which Jake felt was suitably ominous. Acronyms were for government funding, and Bush Jr. III had knocked that on the head after the MORMON simulation results came back. Well, a call from The Simulation was a call you had to answer, and he patched in. The Simulation recounted its calculation paths for the previous evening and then issued a warning: a socioquake, and a biggy. So biggy, infact, that Jake sat back on his bed and considered which family members he might phone before he looked into it. As it was, The Simulation had already booted his home terminal and was demanding a response. Jake nevertheless called his sister in Baltimore and blearily told her answering machine to stock up on tinned food before turning his attention to the terminal cube. The globe floating in the cube showed the North East US throbbing a cancerous red, the devastation sweeping out through both the transport and financial networks to other areas of the world, death tolls accelerating each time Jake rubbed his eyes and tried to focus on the issue.

The Simulation was a construction so beautiful that aged programmers had been known to break down and weep before it**. A total behavioural encapsulation of the human race, gridded to every known dataset the globe collected.** Ok, so that beauty came at a price: $200 billion a year and 20 separate wind-farms; it had also turned Jake from a bright young fella into an insular wreck. But, oh, the wonder of it. The Simulation was still based around its core functionality – socio-economic prediction, but it could do so much more. The ability to **upload behaviours across a variety of social scales** allowed the system to act as a gigantic **validator of social logic** – new discoveries uploaded as viruses by everyone from transactional psychologists to playground monitors could compete in the system for sociological survival. The Gödel Module integrated new behavioural models and **assessed their consistency with current theory and for internal completeness**, identifying strings of behavioural incompatibility, pinpointing potential hypocritical paradoxes, and revealing new holes in our understanding. Once a new behaviour was validated logically, it was tested quantitatively – introduced to a generalized model first, and then to models of increasing socio-economic detail, until finally being **dynamically added** to The Mirror – The Simulation's dark and chaotic identical twin. Only if the Mirror held up to this introduction and could **back-predict the entire data timeline at key spatio-temporal scales based on a variety of supports** would the behaviour be added to The Simulation itself, to be run either in the full model, or one of the Point-of-view models that **slanted the simulations towards a particular world-view**. Environmental and Climatic modules had been made interoperable with The Simulation early on – initially there had been many such models, but funding unification under the UN following the submersion of Florida and the Netherlands had rapidly driven the cream of the crop to solidify into a single entity, held together by **Translation Schema**. By and large, the models were run together these days, though short term weather changes were sometimes still predicted independently.

It wouldn't be inaccurate to describe The Simulation as conscious. Although spoken language was largely redundant in modelling human interactions, any given agent in the system could pass the Turing Test. The general public largely loved The Simulation, or, more specifically, their piece of it, and regularly **updated their profile in the machine with tit-bits of knowledge or additional rulesets**. The key buy-in by the public had come after the development of the "Voices of the Past" project – a Sino-Korean genealogical tool that gave people the nearest to immortality they could hope for. Voices of the Past allowed people to update the agents representing them with the knowledge and rulesets which they lived by, often **automatically generated by the data capture systems they interacted with**. When they passed on, the agent was made available to their descendants for querying – a person could ask their family ancestors any question: "*is he a good man*?"; "*what do you think of the market today*?"; "*how do I cope with the despair*?", and get a variety of trusted answers. People were suddenly able to explicitly and directly tap into the huge bank of experience that had, implicitly, made them who they are. Initial the Voices were generic and thin, the ghosts of personalities from the past, but with each iteration the knowledge capture became more sophisticated and the recently deceased gave enchantingly personality-based advice. Most importantly, the service was free to those who uploaded. The system cross-compared everyone, calculating which knowledge was personal and which more general, building up a **core map of human intelligence** as well as a hive intelligence constructed from the world's 4.3 billion citizens. In addition, the system was globally self-aware – iteratively **predicting its own influence**, both socio-economically and environmentally – the latter dependent on the workload in the 20 or so processor farms dotted around the globe. In the early days they'd had given The Simulation a voice constructed by parsing Slim Pickens' dialogue from Dr Strangelove – the kind of geek humour they had time for back then – but funders had found it too disturbing; now, if The Simulation spoke at all, it was as James Earl Jones. But the main interface was haptic, via the cube – the very cube Jake was sat in front of now.

The socioquake prediction was for three days time. Jake checked it wasn't a policy-experiment, but confirmed his fears that it was from a work-a-day run, predicated only on the current state of society – a state in turn garnered from every dataset people meandered through: from CCTV estimates of eating habits, to social-nets parsed from phone calls, to the taxation trackers in the last few remaining private cars. Jake grabbed the globe and drilled down into the North East. It was a mess. The three day state was one of total social collapse. Not simply the usual odd pocket of riots or demonstration, but a mélange of armed gangs, ethnic cleansing, and spontaneous group murders across the whole social spectrum. Jake stared sweating at the globe in a moment's paralysis, before the spreading red pulled him together. There had always been the potential that society would flip away from its stable attractors, but the homeostatic forces were immense – Jake never expected it to bleed from theory to his terminal.

Lost for a strategy to deal with the unfolding horror, Jake resorted to the standard techniques. Cancelling the press forecasts, he first pulled up the **error surface associated with the prediction** – while it was fluctuating wildly for Europe, the errors were tightly dampened for the US, in particular the Eastern Seaboard. He then sliced into the errors to reveal the major contributors, first for the current prediction, then the most solid contributors over the past three days, probing the time periods and networks those errors had acted over. He saw no fluctuations in the errors to make him doubt the predictions. By the time he finished, three hours had passed – too long, and he knew it; even The Simulation had limits. If he were to prevent the quake he needed to do two runs: an investigation and a preventative-policy estimation; now he only had time for one and a half. Damn; it was a coin-flipper for sure. Gambling that an investigation might turn up an obvious cause, while a policy run might just stall things for 24 hours, he opted for a full investigation and a low-grade policy suggestion. Setting the run going, he grabbed the globe again, and began to trace back the emergent properties. Prior to the run, he only had maps of data flows **the system had tagged as interesting and unusual**. These trends were good for tracking memes, but for divisible commodities they often appeared and then vanished as individuals dispersed materials to larger groups. Despite this, there did seem to be a spatial autocorrelation in, of all the crazy things, *bean purchases* in areas that later generated problems. The system hadn't flagged the origin of this, and Jake bemarked it for later mining.

An hour later the investigative run was complete, and the low-grade policy prediction began, searching through multiple policies that might **kick the system back into a stable state by synchronizing the various distribution networks**. The right kick, at the right time, might just rescue it. But there was limited time to run real simulations, and the low grade run used some pretty rough heuristics. While The Simulation ran its merry way, Jake started on the investigative key statistics, tracing them back through the spatio-temporal maps using the **probabilities of causality** generated by the investigative run. As time charged on, Jake became more intense, tearing off networks, flipped back and forward through probability differentials, and running small causality calculations on network loops that largely dissipated to nothing. All was apparently to vain, but still he cut back through the simulation. Finally he noticed a zone of spending: panicked buying preceding the wave of violence some five hours before the real trouble broke out. He followed it forward, spreading out across the East, and then he flipped back, watching the wave shrink towards its origin. The wave propagated back, back and back from the Seaboard to the North East, and then a sudden turn, south to Maryland, and then…the screen froze. Crap, The Simulation had reached the current moment – the Critical Horizon, Jake slumped back into his chair. It was too late. Whatever it was, it was already happening. Jake stared at the screen, only to be slapped back into awareness by a broad voice indicated the policy run had been completed. One last hope. The result, with all the desert-dry wisdom of the Delphic Oracle, was a taciturn "*Three thirty three; buy no beans*".

Jake gave the cube a look that might have been acceptance, or resignation, or the dull flash of years of calculated ennui imploding to nihilism, and grudgingly pushed himself up. He walked to the window, and let the morning sun stream warm across the room. Pulling on a pair of antiqued jogging pants and trainers, he walked out of his door and across the plaza. Half way across the square he stopped mid-step, backtracked to his block, and stuck a $200 bill into the cup of an old critical geographer who had resorted to begging from his doorway, then, turning unhurried into the sun once more, he set off across the plaza and down the strip towards the local gun mart, reluctantly fingering the credit card in his pocket.

INDIANA UNIVERSITY

Department of Geography

701 E. Kirkwood Ave.
Student Building 120
Indiana University
Bloomington, IN
47405-7100

Michael Goodchild
Department of Geography
University of California – Santa Barbara

Date: January 29, 2007
Re: Agent-based Modeling Workshop

Dear Dr. Goodchild:

This is an application to the recently announced Workshop on Agent-Based Modeling of Complex Spatial Systems. The following describes my views on the direction of this field, my prior research related to agent-based modeling, and my future planned directions.

I have been working with agent-based models with respect to land cover change research since 2000 with a particular emphasis on household decision-making dynamics and landscape outcomes. Since this time I have seen agent-based models move from relatively abstract representations to those more tightly linked to empirical foundations. Yet, while those employing agent-based models of land cover change (as an example) often use observed land cover data to 'validate' the model (avoiding the discourse on what 'validation' means for the moment), I believe there has been less attention given to validate the unique characteristics of agent-based models that make them more attractive than other modeling approaches. In my mind, these characteristics include: 1) the ability to represent agents and their decision-making strategies heterogeneously and 2) the ability to explicitly incorporate interactions between agents.

Now that the agent-based modeling community has made progress in supporting models with empirical data, I believe that the next logical step is to test the performance of our models with data that explicity tests the above two characteristics. This is admittedly a considerable challenge as the data collection costs necessary is potentially prohibitive for many projects. To truly assess the role of agent interactions in a system will require complex new datasets to be collected that are both rich and longitudinal – quite a daunting task. However, this next step should be considered if we are to convince the broader community (both modelers and others) why agent-based models are more suitable for some tasks than say spatial regression or cellular automata.

My personal background in agent-based modeling began with a a NSF award from the Biocomplexity program for a project titled: "Biocomplexity in Linked Bioecological-Human Systems: Agent-Based Models of Land-Use Decisions and Emergent Land-Use Patterns in Forested Regions of the American Midwest and the Brazilian Amazon" on which I was a Co-PI.

From this research I have published manuscripts utilizing agent-based modeling in *International Journal of Geographic Information Science*, *Environmental Management*, and several book chapters along with manuscripts in review with *Geoforum*, *Land Change Science,* and *Journal of Economic Dynamics and Control.* These papers have broadly explored topics including the role of scale dependence in agent-based models, the use of agent-based models for backcasting, and the integration of methods from experimental economics and agent-based modeling.

Other activities include co-organizing (with Steve Manson) a special issue of the journal *Environment and Planning B* focused on modeling and complexity in geographic research. This special issue is the product of a series of organized sessions on geographic complexity at the 2005 AAG meeting. The special issue is planned for publication in the March 2007 issue of EPB.

My most recent research involves the use of agent-based modeling to explore the dynamics of reforestation in Indiana and Sao Paulo, Brazil. I am the principal investigator of a new project titled “Dynamics of Reforestation in Coupled Social-Ecological Systems: Modeling Land-Use Decision Making and Policy Impacts” recently funded by the NSF HSD program. In this extension of previous research, we will incorporate a more diverse set of agents to explore land cover change dynamics. In particular, we will represent actors such as NGO’s and governmental officials and their interactions with household level actors in these new modeling efforts. This research will also be tightly integrated with complex physical models (hydrology, forest ecology) at various spatial scales of analysis. As on prior research, the approach of this project is highly multi-disciplinary with colleagues from anthropology, hydrology, forestry and political science.

While I expect this to be a popular workshop, hopefully there is room to allow me to attend. I look forward to hearing if that is the case.

Sincerely,

Tom P. Evans

Philosophical and practical limits of Agent Based Models as viable systems for discovering and verifying new geographical knowledge

Mark Gahegan

GeoVISTA Center
Department of Geography,
The Pennsylvania State University, USA.

David O'Sullivan

School of Geography, Geology and
Environmental Science,
University of Auckland, New Zealand.

New analysis and modeling approaches often call for reconsideration of methodological and philosophical stances if they are to be used appropriately. This is particularly true of Agent-Based Models (ABM), which can play a wide range of roles in scientific investigations, from characterizing emergent properties of data through prediction of future states, to positing explanations of possible causal mechanisms. In the biological simulation community, a similar recognition is summed up by Peck (2005) thus: "*Philosophers and practitioners of science are recognizing that simulation models are a new kind of tool that defies the categories, uses and restrictions found in the historical use of mathematical models*". Such models do not sit easily with the traditional view of models and their roles in geographical analysis (e.g. Chorley, 1964). While we do not agree that simulation is a 'new kind of science' we do believe that its application and interpretation, and legitimate roles and limitations, are not yet well understood. Equally, real-world experiments are not practical for many kinds of broad-scale, geographical inquiry, and simulation models allow exploration of alternative realities and responses to change. Thus, we must learn to use simulation technology in effective and defensible ways.

How do we validate such complex models? There is a danger that simulation models can become self-fulfilling prophesies because they often conflate different analysis activities that would ordinarily stand alone and be *independently* scrutinized. For example: data collection, model synthesis, numerical analysis, validation and presentation (e.g. Gahegan & Brodaric, 2002) are traditionally disjoint activities where uncertainties at each stage are accessible for independent investigation. In simulations, some (or all) of these activities may become intertwined. For example, data may be imputed, loaded into individual agents, which then interact via a (possibly evolving) set of rules to produce outcomes that appear realistic or useful. But mere plausibility is no guarantee that the explanations derived are true in the world. If the modeler expects certain outcomes, constructs rules and gathers data accordingly, then tests various models until their behavior matches expectations, then there is little independence and a lot of bias. Over-fitting will be rife and there is little chance for novel or unexpected outcomes to emerge.

How do we know what is going on 'inside' simulation models? Visualization in support of ABMs remains quite primitive, and typically unable to provide much insight into the complex interactions and states of many independent actors. Interactions become so complex that it is not possible to be sure exactly how the model is behaving. We may be confident in model outcomes as a realistic representation, but causal mechanisms (and explanations) may remain elusive. A related concern is that difficulty in observing what is going on 'inside' models often means that they are viewed in terms of (pre-selected) aggregate measures that reinforce the modeler biases.

Also, disparities between the scale at which models are developed (individual decision-making) and evaluated (aggregate outcomes) raise further questions about validation processes.

How geographical are ABMs, really? ABMs appear to be explicitly spatial, but that does not mean that they are inherently geographical. For the most part, ABMs represent only a very impoverished, grid-based geography, with poor handling of boundary conditions and without any inherent structure within the space (Gilbert & Bankes 2002). There are only a few exceptions, (e.g. O'Sullivan et al, 2003; Brown et al. 2005). That they are often considered as collections of actors within a geographical space does not mean that they magically resolve or avoid any of the classical statistical pitfalls that beset geographical analysis. In particular, it is unusual for models to reflect the multiple geographical (and social) scales at which decisions are made. ABMs are at their most convincing in contexts where constrained actors make decisions in 'reactive' ways, the prime example being various models of pedestrian behaviour (see e.g., Helbing et al. 2001), but such contexts represent a small percentage of cases where human activities make a difference.

How do we avoid merely reflecting our own biases? Arising from all of these questions, and as Bankes (2002) suggests there is a tendency to focus on those aspects for which straightforward behaviors can be constructed, or for which good data are available, but to ignore or play down other more problematic aspects. This leads to bias in the results (in terms of both explanation and prediction) and points to an important and elusive question: "*Did we really represent and explore the space of all plausible models?*" Currently, the search for a solution tends to stop when a useful model is produced (perhaps tested by goodness of fit to some desired outcome, or more informally because its behavior 'seems right'). This question does not apply to simpler forms of predictive modeling where the outcome can be validated straightforwardly (e.g. predicting stream discharge). But if the outcome is complex, the data uncertain, or the aim is to create an explanatory model (e.g. predicting a landcover change surface) it is very likely that a family of solutions exists, all of which would perform equally well—within the wide margins of confidence (*equifinality*). The internal differences exhibited by a family of models that produce similar outcomes could tell us much about the nature of the systems we analyze, including their stability and the confidence we should place on predicting future states. **Representing the hypothesis space of solutions and the regions within it that a simulation has (and has not) explored is a very difficult but important problem that needs to be addressed. Likewise highlighting and comparing the parameterization of solutions that produce similar outcomes. (See figure at end).**

Some issues we are interested in and on which we can most readily contribute:

1. Providing ways to better understand—by visualization—the detailed inner working of ABMs
2. Building ABMs that are more geographically explicit
3. Developing measures of confidence in the results obtained from ABMs, for example by representing the path that potential model solutions have taken through the hypothesis space.
4. Adding to the theoretical understanding, and creating associated guidelines on best practices for the use and evaluation of ABMs in geographical analysis.

References

Brown, D. G., Riolo, R., Robinson, D. T., North, M. Rand, W. (2005). Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems* **7**(1), 25-47.

Chorley R. J. (1964) Geography and analogue theory. *Annals of the Association of American Geographers*, **54**(1),127-137.

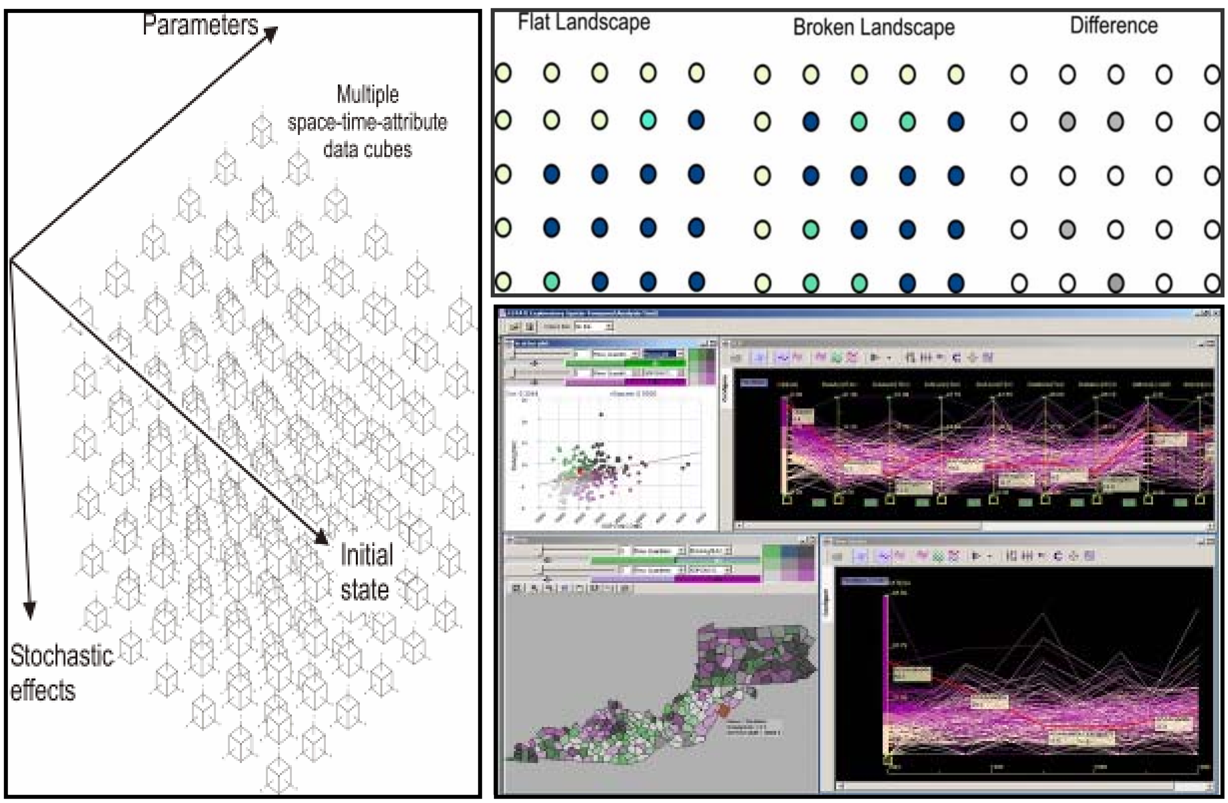
Gahegan, M. and Brodaric, B. (2002) Computational and Visual Support for Geographical Knowledge Construction: Filling in the gaps between Exploration and Explanation. In D. Richardson and P. van Oosterom (eds.), *Advances in Spatial Data Handling, 10th International Symposium on Spatial Data Handling*, pp. 11-26 (Springer, New York).

Gilbert, N. and Bankes, S. C. (2002). Platforms and methods for agent-based modeling. *Proceedings of the National Academy of Sciences of the United States of America*, **99** (Supp. 3), pp. 7197-98.

Helbing, D., Molnár, P., Farkas, I. J. and Bolay, K. (2001). Self-organizing pedestrian movement. *Environment and Planning B: Planning & Design*, **28**(3) 361-83.

O'Sullivan, D., Macgill, J. R. and Yu, C. (2003). Agent-based residential segregation: a hierarchically structured spatial model. In C. Macal, M. North and D. Sallach (eds), *Agent 2003: Challenges in Social Simulation*, pp. 493–507 (Argonne National Laboratory: Chicago).

Peck, S. L. (2004) Simulation as experiment: a philosophical reassessment for biological modeling. *Trends in ecology and evolution*, **19**(10), 530-534.



LEFT: Visualizing the space of *possible* models, with the aim of uncovering *plausible* models—then examining their similarities and differences. TOP RIGHT: Examining the effects of changing parameters on model outcomes (left & middle grid), explicitly representing where different outcomes occur (right grid). BOTTOM RIGHT: Visualizing the parameters and outcomes in sets of model runs.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

The complexity of agent-based spatial models

Nigel Gilbert
Centre for Research in Social Simulation, University of Surrey

In the manuscript for a text that I have just sent to the publishers (Gilbert, 2008), I distinguish between abstract, middle-range and facsimile models (Boero & Squazzoni, 2005) and propose that the requirements for validation differ between these types:

> **Abstract models** aim to demonstrate some basic social process that may lie behind many areas of social life. A good example is Epstein and Axtell's pioneering book on *Growing Artificial Societies* (Epstein *et al.*, 1996), which presents a series of successively more complex models of the economics of an artificial society. … With these models, there is no intention to model any particular empirical case and for some models it may be difficult to find any close connection with observable data at all. … How then might such models be validated? The answer is to see such models as part of the process of development of theory, and to apply to them the criteria normally applied to evaluating theory. That is, abstract models need to yield patterns at the macro level that are expected and interpretable, to be based on plausible micro-level agent behavioral rules, and, most importantly, to be capable of generating further, more specific or 'middle-range' theories (Merton, 1968).
>
> **Middle-range models** aim to describe the characteristics of a particular social phenomenon, but in a sufficiently general way that their conclusions can be applied widely to, for example, most industrial districts rather than to just one. The generic nature of such models means that it is not usually possible to compare their behavior exactly with any particular observable instance. Instead, one expects to be satisfied with qualitative resemblances. This means that the dynamics of the model should be similar to the observed dynamics and that the results of the simulation should reveal the same or similar 'statistical signatures' as observed in the real world, that is, the distributions of outcomes should be similar in shape (Moss, 2002).
>
> **Facsimile models** are intended to provide a reproduction of some specific target phenomenon as exactly as possible, often with the intention of using it to make a prediction of the target's future state, or to predict what will happen if some policy or regulation is changed. For example, a business may be interested in finding the consequences for their inventory level of reducing the interval between sending out restocking orders. It is likely to require a model that precisely represents all their suppliers, the goods each supplies, and the unit quantities of those goods in order to be able to make reasonable predictions.

> Another, very different example, is the work by Dean *et al.* (1999) on the Anasazi Indians in South West United States. These people began maize cultivation in the Long House Valley in about 1800 BC, but abandoned the area 3000 years later. Dean *et al*.'s model aimed to *retrodict* the patterns of settlement in the Valley and match this against the archaeological record, household by household.
>
> If such exact matches can be obtained they would be very useful, not only as a powerful confirmation of the theory on which the model is based, but also for making plausible predictions. However, … Most social simulations contain some element of randomness. For example, the agents may have initial characteristics that are assigned from a random distribution. If the agents interact, their interaction partners may be selected randomly, and so on. The same is presumably true of the social world: there is a degree of random chance in what happens. The effect of this is that running the model a number of times will yield different results each time. Even if the results are only slightly different, the best one can hope for is that the most frequent outcome – the mode of the *distribution* of outputs from the model – corresponds to what is actually observed (Axelrod, 1997; Moss, 2002). If it does not, one might wonder whether this is because the particular combination of random events that occurred in the real world is an outlier, and that if it were possible to 'rerun' the real world several times, the most common outcome would more closely resemble the outcome seen in the model!

I think it is true to say that the majority of agent-based models of complex spatial systems fall into the second and third of these categories, that is, they aim to be either middle-range or facsimile models. In the workshop, I want to explore the relevance of the first category and consider what it might mean to have abstract spatially explicit models. Among the implications are:

a. More effort might be put into developing ABMs representing the classic theories of geography. Issues for discussion include:
    1. Which theories?
    2. Is it possible and useful to cast them as ABM?
    3. What progress has already been made?
    4. What benefits might there be in doing this?
b. The abstract models approach implies a certain kind of epistemology: a realist perspective that assumes that there are 'real', generative social processes (Epstein, 2007) that yield the observable features, and that these social processes are recurrent in different social phenomena (thus, what I have called 'abstract social processes'). Examples are:
    1. Markets. While the observable characteristics of markets vary widely, there is some basic, recurrent and abstract underlying social processes from which market institutions emerge;
    2. Cooperation in the face of social dilemmas. There are many situations in which the long-term socially optimum behaviour is not the individually optimum one, but where cooperative behaviour has 'evolved'. There is now a rather large literature on how this works.

Is this epistemology defensible, and how can it be made more precise?

c. What methodological advice can be offered about how one should develop abstract models? How does one validate an abstract model?

References

Axelrod, Robert. (1997). Advancing the art of simulation in the social sciences. In R. Conte, R. Hegselmann & P. Terna (Eds.), *Simulating Social Phenomena* (pp. 21-40). Berlin: Springer.

Boero, Riccardo, & Squazzoni, Flaminio. (2005). Does Empirical Embeddedness Matter? Methodological Issues on Agent Based Models for Analytical Social Science. *Journal of Artificial Societies and Social Simulation, 8*(4) http://jasss.soc.surrey.ac.uk/8/4/6.html

Dean, Jeffrey S, Gumerman, George J, Epstein, Joshua M., Axtell, Robert L., Swedland, Alan C., Parker, Miles T., et al. (1999). Understanding Anasazi Culture Change Through Agent-based Modeling. In T. Kohler & G. Gumerman (Eds.), *Dynamics in Human and Primate Societies*. Oxford: Oxford University Press.

Epstein, Joshua M. (2007). *Generative Social Science: Studies in Agent-Based Computational Modeling*: Princeton University Press.

Epstein, Joshua M., Axtell, Robert, & Project. (1996). *Growing artificial societies : social science from the bottom up*. Washington, D.C. ; Cambridge, Mass. ; London: Brookings Institution Press : MIT Press.

Gilbert, Nigel. (2008). *Agent-based models*. Thousand Oaks, CA: SAGE.

Merton, Robert King. (1968). *Social theory and social structure* (1968 enl. ed.). New York: Free Press.

Moss, Scott. (2002). Policy analysis from first principles. *PNAS*, *99*(Suppl. 3), 7267-7274 http://www.pnas.org/cgi/content/abstract/99/suppl_3/7267

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Evaluating Agent-Based Spatial Models

Rich Harris
School of Geographical Sciences and Centre for Market and Public Organisation, University of Bristol

My interest in Agent-Based Spatial Models is, in truth, driven less by e-science / e-social science *per se*, and more in an on-going curiosity in the extent to which social 'systems' can or should be modelled, thence predicted and/or 'explained' using quantitative and/or computational methods. This curiosity is better expressed by Peter Haggett who wrote, in 1994, that:

> there may be limits to the predictability of human behaviour which makes prediction in the social sciences fundamentally flawed … But, as one who finds continual refreshment from the work of colleagues in physical geography, I consider that if the boundary exists it should continue to be actively probed (Haggett, 1993: 18)[1]

In short, then, I am interested in what agent-based spatial models can contribute to developing socio-spatial theory: what can they usefully tell us about socio-economic systems that was not known already? Alternatively, if the primary purpose of agent-based spatial models is less to extend theory, than to test it, then my interest is in identifying real-world applications and genuine problem-solving for these, as for other geocomputational toolkits (noting, in particular, Martin's 2005 interest in the potential of e-social science to support the development of geocomputation).[2]

If my 'Bristol upbringing' explains my interest in the possibilities of computational social-science then it also accounts for my caution – exposed, as I am, to the waves of postmodernism, post-structuralism, 'non representational' theory and other epistemological and ontological turns away from the quantitative/computational. Whilst it may be simplistic to suggest that contemporary social science is characterised by the triad of (1) the theory-led and deductive models of e.g. economics, of (2) the data-based, inductive but scientific and mathematically-informed methods of statistics and (3) a rejection of structured forms of enquiry and explanation in, for example, some areas of human geography, still I would argue that many methods of geocomputation and computational e-social science sit away from any clearly identifiable 'camp' and that

[1] Prediction and Predictability in Geographical Systems. Transactions of the Institute of British Geographers, 19, 6-20.
[2] Socioeconomic GeoComputation and E-Social Science. Transactions in GIS, 9, 1-3.

may underpin their credibility or potential for uptake (a point that Couclelis made in relation to geocomputation back in 1998).[3]

To put this all another way, what sort of enquiry (knowledge formation) do agent-based spatial models support, or are governed by? Do they already or can they be shown to have their place within socio-economic research? Do they need to conform to more orthodox traditions or to be defended against philosophical fashions?

A third way of looking at this is to ask whether I could demonstrate, to my students of Derrida, Foucault, Deleuze and Guattari, the value of agent-based spatial models, and convince them of it.? There are two issues, here. One is of a 'flagship' model that showcases potential and raises interest (it probably exists; my ignorance of such matters is not in question here!). The second concerns usability. This becomes increasingly important if we consider the potential to mount such models on, for example, the UK's National Grid Service (NGS): a computational grid of high performance machines primarily developed under the UK's e-science research funding.[4] From experience – and having been on a (so-called) training course – the NGS is in its infancy and extremely difficult to use. The computational power is certainly there but using it is far from straightforward.

To summarise, I am interested in answering the following questions and thence convincing sceptical colleagues:

- What have agent-based spatial models got to do with social science and, in particular, human geography?
- What disciplinary traditions are they founded on, how much so, and are the purposes of the models more for developing or for validating theory (then, how is this done? Can it be done?[5])?
- How should I use them and what are some of the pedagogic considerations of teaching with/about them?
- How can the potential of the UK's National Grid Service be harnessed for agent-based spatial models and other spatial, computational modelling?

Essentially this all amounts to re-stating the aim of the ESRC's National Centre for e-social science: "to investigate how innovative and powerful computer-based infrastructure and tools developed over the past five years under the UK e-Science programme can benefit the social science research community."[6]

[3] In Geocomputation: a primer (eds. Longley, Brooks, McDonnell & Macmillan), Chichester: Wiley, pp. 17-29.
[4] www.rcuk.ac.uk/escience
[5] What does prediction mean, for example, in the context of an open and changing system (which is not actually singular)?
[6] http://www.ncess.ac.uk

Modeling Motion Relations for Moving Objects on Road Networks

Kathleen Stewart Hornsby
Department of Spatial Information Science and Engineering
National Center for Geographic Information and Analysis
University of Maine
Orono, ME 04469-5711 USA

Introduction

I have been working on topics relating to spatio-temporal data modeling for some time. One of these topics, relates particularly well with the focus of this workshop, *Agent-Based Modeling of Complex Spatial Systems*. With this research, I am interested in modeling the semantics associated with different types of moving entities, in particular, the semantics of moving vehicles on road networks, and the corresponding extensions to data models that are necessary for next-generation information systems to support vehicle navigation more fully. The movement semantics that we are examining relate to the position of one moving vehicle relative to another on the road network. These *motion relations*, based on the relative positions of pairs of moving vehicles on roads, capture the kinds of cognitive semantics that are especially meaningful to drivers and other active participants (e.g., bicyclists) as they travel and interact with each other on roads. For example, *is the car behind my vehicle a police car? Is that an ambulance coming towards my vehicle in the other lane? What kinds of vehicles are currently in the vicinity of Maple Street and 4th Avenue?*

For this work, it is assumed that vehicle positional data are collected from a geosensor network and stored in a spatial database. These datasets are used to compute a set of basic motion relations including *isBehind*, *driveBeside*, and *passBy*. The motion relations describe the positions of one vehicle relative to another in lanes of traffic and serve as a foundation for representing *individualized* perspectives of dynamic transportation networks. These perspectives capture details of movement of entities (vehicles) modeled as objects, for a user-defined spatio-temporal region of interest in the transportation domain. For example, for any given region of interest and an interval of time, it is possible to derive the kinds of moving objects and their corresponding relations to each other with respect to a reference object, affording a refined perspective of the kinds of moving objects currently in the vicinity of that object. These semantics are important for understanding, modeling, and querying the behavior of moving entities in a modeled road network, as well as for annotation and enhanced indexing.

Moving object data from geosensor networks

For this work, a fixed-length linear referencing model is used in conjunction with a point-location approach to represent traffic movement data. The network of sensors is distributed along a roadway so that each one observes a specific number of fixed reference positions. We assume that at any given time *t*, only one object can be at a specific reference position for any lane (e.g., one object can not be on top of another object). Data collection begins when movement is detected. From the initial point of movement, a series of sensor readings are collected at a fixed time from one another. The positions of each individual object within the range of the sensor (relative to a reference point) are then stored for each reading. Each moving object is assigned a unique identifier, and the linear extent of the moving object and its midpoint are stored as well. Based on this framework, a database representation is developed to provide the basis for supporting queries that describe different kinds of motion relations. A relation *SensorDat* with attributes, *objectID*, *sensorID*, *laneID*, *position*, and *time* is defined for storing location readings within the sensor from the network. Details of the moving objects, on the other hand, are stored in relation *ObjData* with attributes, *objectId*, *class* and *length*. The attribute *class* corresponds to object classes that model the kinds of vehicles moving on the road network.

Basic motion relations for moving objects

The basic actions of two or more moving vehicles on a road form the foundation for a typology that distinguishes a set of basic *motion relations* (i.e., an elementary set of relations between two moving objects). These movement types are restricted to movements that result in a change of location. The basic motion relations that are the focus of this research are *isBehind(A,B,T)* and its converse relation, *inFrontOf(A,B,T)*, *driveBeside(A, B,T)* and *passBy(A,B,T)* (Figure 1). The relation *isBehind(A,B,T)* and its converse relation, *inFrontOf(A,B,T)* describe the relative spatial relation between two moving objects *A* and *B* (e.g., land vehicles) in the same lane of traffic at time *T*, such that no other object is between *A* and *B*. The case where one vehicle goes by another in an adjacent lane of traffic while traveling in the same direction at time *T*, and assuming no vehicle is in between, is captured by *driveBeside(A,B,T).* The alternative case of two vehicles moving in adjacent lanes of traffic while traveling in opposite directions at time *T*, is referred to as *passBy(A,B,T)*. The *driveBeside* and *passBy* relations are both symmetric.

The role of agents

At present, we have been working on the database aspects of this research, formulating queries in SQL that compute the motion relations for a given set of vehicles. However, it is very interesting to consider how agent-based approaches could be applied to this work. For example, each vehicle could have an embedded sensor that allows it to receive communications from the road sensors. These communications would inform the agents of the position of the vehicle on the road network. In addition, the onboard sensors would allow for communication with other vehicles that are around it on the road network. In this setting, the data streams for the agents would come from within the vehicles themselves as well as from the sensors in the transportation network. And in this way, vehicles would be informed by agents about the position and types of vehicles around them, allowing them to move into the lane of traffic that will facilitate their progress, or communicate with other vehicles such that they change position with each other. Based on this, one could foresee scenarios where, for example, service vehicles (e.g., taxis) could communicate with each other with respect to their position in the network via agents, and move to optimize their travel routes.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Revealing Hidden Dynamics in Spatial Data

Alison Heppenstall, CSAP, Leeds, UK

Over the last 15 years, geography has witnessed an explosion in the provision of both computational power and digital spatial datasets. This has brought a greater awareness to geographers of the nature and significance of small-scale individual-level dynamics and their effect on larger scale complex system dynamics. Geographers are now beginning to use concepts and notions from complexity theory to reappraise geographical systems. This marks a significant departure from the traditional treatment of complexity - aggregation of people to groups, and the accompanying statistical treatment of these groups, by empirical techniques for example, regression based models. Such models are now recognised as lacking the ability to detail the effects of small-scale and individual level histories, interactions, or even, in any realistic sense, behaviour.

Advancement in both computing and understanding has been accompanied by the development of new techniques, such as agent-based models (ABM) and microsimulation, for simulating complex systems. These methods have provided us with the ability to begin modelling and analysing the impact of individuals and their behaviour over both space and time. The current range of ABM applications in the literature demonstrates its vast potential as a tool for dissecting and understanding the inner workings of complex systems. One significant advantage of ABM is its ability to produce good simulations without detailed data; unlike our more fortunate colleagues in the physical sciences, social scientists are not blessed with detailed temporal and spatial data. However, I feel that more work is required on developing techniques for extracting additional information (e.g. system structure and behaviour) from available data to enhance our models.

In most ABM applications, information fed into agents is typically drawn from observational or simple empirical analysis. In some cases this may be sufficient, but it is inadequate if our goal is to understand structure and dynamics, patterns and behaviour, how individuals or behaviour evolves and adapts in our systems. We simply cannot capture the underlying structure and dynamics of the environment that the agents are embodied within. To achieve this we need to be able to unravel system dynamics and behaviour of the components of the system, in both real and model data.

Investigations into the applicability of methodologies from other disciplines such as physics, chemistry or mathematics, for analysing and modelling complexity in geographical systems are largely absent. This appears to be a significant omission if we are to fulfil the potential of ABM for examining the impact of small-scale individual dynamics on the larger system. I feel that we need to be concentrating on (i) developing better techniques for analysing real complex systems (to provide more detailed and realistic inputs into ABM) and (ii) developing sophisticated methodologies to analyse the results from ABM.

I believe that this can be achieved if research is directed towards the development of tools for identifying/visualising structure and underlying dynamics/relationships within spatio-temporal data. This could lead us to characterise the behaviour of both individuals within a system and the entire system. Through this type of work, we can begin to (i) identify the many threads that exist within the systems and how these contribute to the behaviour of our systems and (ii) understand which techniques are the most appropriate for our analysis.

For characterising individuals and systems, there are several methods that we can borrow from mathematics and physics. For example phase diagrams allow us to chart the behaviour of one or many individuals or the whole system. This can give us insight into the whether behaviour is stable, chaotic or entirely random. Experimentation using ABM and the notions of, for example, catastrophe theory and bifurcations may also yield useful information. These types of nonlinear behaviour are not produced by “traditional” models meaning that these issues have not, to date, been widely addressed in geography.

Wavelets allow us to decompose the system into signals – further work in this area may help to identify “noise” or weak signals in data; we have no way of either identifying or knowing whether these weaker noisy signals over a period of time stabilise and control systems. Investigation into other methods such as recurrence plots may help us to visualise hidden structures and dynamics within data. The results of this type of work may help us to characterise our systems and their behaviour.

The areas briefly alluded to above all have great potential for our understanding by providing new information and insights into our systems that we can feed into our models. However, considerable work needs to be done to develop these methods for application to spatio-temporal systems. Many techniques available in other disciplines have been specifically developed for time series or spatial data (for example image processing). In geography, the spatio-temporal nature of the data adds additional problems. Available techniques require modification and development for geographical applications. I would like part of the workshop to concentrate on this area.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Andy Hudson-Smith
CASA, UCL

Agents in Virtual Worlds

My work is focussed on generating a variety of graphic media for various projects in CASA ranging from our 3D GIS model of Greater London to panoramic imaging of urban environments through to real time monitoring of local urban climates. Much of this work is deigned for dissemination to both professionals and less expert community groups through primarily web based services. A variety of my work is shown on my blog http://www.digitalurban.blogspot.com/ but here I will focus on the work I have been doing with virtual worlds and the representation of presence and agent behaviour in digital environments.

Agent based modelling is traditionally a 2D discipline in academia which has an emphasis on pedestrian and transport simulation within the field of spatial analysis. Yet advanced agent based modelling is possible, and indeed common, within 3D multi-user environments. We examine the use of these agents, based on our own experience, detailing the use of agents in virtual worlds and gaming environments.

We first examine the use of bots in ActiveWorlds, a multi-user environment whereby the user is represented as an avatar (Figure 1); this is a common theme throughout such systems.

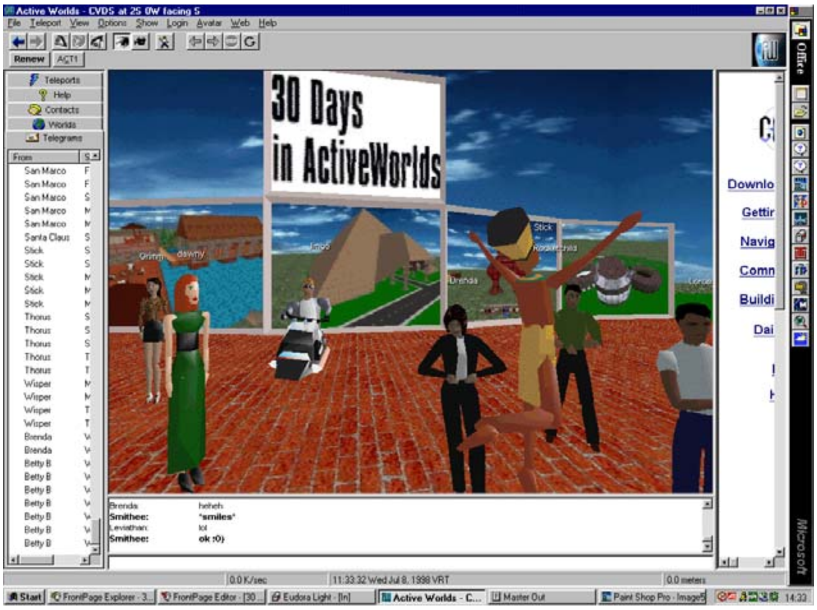


Figure 1 – Avatars and Agents in ActiveWorlds

Agents can be inserted, with varying degrees of complexity and autonomy, into the ActiveWorlds environment. We have examined the use of such agents, detailing the ability to relate to shortest path analysis, object construction, artificial conversation according to proximity, and spawning behaviour.

The use of such agents in multi-user environments is often controversial and the use of Non-Player Avatars or NPC's, as agents are more commonly known, has been banned in many systems. Of note is the ability of agents to build and interact with their environment combined with the ability to clone not only objects but also themselves through self spawning. This has led to a number of examples of virtual world vandalism carried out by autonomous agents, as indeed we note in our 30 Days in ActiveWorlds paper.

There are calls for a ban on NPC's in the environment known as Second Life. Of note is the 'CopyBot' an agent which can clone any other object, this is significant in terms of Second Life as its economic system is based on the purchasing of objects to use in the environment – such as a new house or car. CopyBot can be released into the world and simply clone any object, including avatars, with additional scripting it is then been possible to respawn that object allowing users identity to be stolen and then to be surrounded by multiple copies of themselves. In essence this is the base of crowd simulation yet a crowd which contains a mix of agents and genuine avatars.

Agents in virtual worlds are therefore controversial, yet one only has to look towards the more powerful gaming environments to see where the true development of agents is taking place. The most commonly known game with agents is 'The Sim's' where you have the ability to toggle 'Free will' on and off. Using this function it is easy to understand the complexity of agent systems within game engines.

A number of these games ship with their own 'Sandbox' mode allowing the game engine to be edited and user content added, normally through 3D modelling software such as 3DMax. However, it is not only the 3D section of the games that can be edited, you can also change and add agents opening up the possibilities to agent based modelling in custom made 3D environments. We have examined the possibilities of such agent based modelling within the 'Oblivion' game engine and the results are to date encouraging.

References

Schroeder, R., Huxor, A., **Smith, A**. (2001) Activeworlds: Geography and Social Interaction in Virtual Reality, *Futures*, 33, 569-587.

Smith A. (2002) 30 Days in ActiveWorlds: Community, Design and Terrorism in a Virtual World, in *The Social Life of Avatars, Presence and Interaction in Shared Virtual Environments* (Ed) Ralph Shroeder, Springer, 77-89.

Agent-Based Modeling Position Statement

Kevin M. Johnston and David Maguire (ESRI)

This position paper is written from the perspective of ESRI as a developer of commercial-off-the-shelf GIS products that ship in volume. Our overarching goal is to build software that solves important problems in a widely applicable way. As such, solutions must be robust, reliable, generic, deliver acceptable performance on large, real world data sets, and be easy to implement.

ESRI has 25 years of experience creating spatial modeling tools for vector and raster analysis. In the last five years we have built a geoprocessing modeling environment that synthesizes our experience in spatial modeling into an easy to use, flexible framework. Within the framework hundreds of spatial modeling functions are available each as individual tools that can be accessed from dialogs, the command line, a graphical modeling environment (ModelBuilder), scripting, or from any programming language. The framework is fully object–oriented and is built on .Net standards. In using ModelBuilder a model is comprised of data that are connected to spatial and aspatial tools to create a process to perform some function. Processes can be linked to one another to create simple as well as complex models. This framework has been used extensively to create static, cartographic process models that encompass classic map analysis/algebra operations.

In the latest 9.2 release of ArcGIS we expanded the geoprocessing environment, in particular ModelBuilder, with looping capabilities and the ability to model stochastic events through a series of random number and data creation functions. With these new capabilities we have explored how to add time to models and how to create various simulation scenarios in a dynamic modeling environment. We have been particularly interested in how to develop process models (e.g., a fire growth model), process models with stochastic events (e.g., a fire growth model with spotting), error analysis through simulations (e.g., examine the effects of error in a DEM on a stream delineation model), and sensitivity analysis (e.g., how does the output change with a 5% change in the distance to roads in a housing suitability model).

With a strong interest in agent-based modeling, we worked with Argonne National Laboratories, the creators of the RePast ABM system, to build Agent Analyst. Agent Analyst is a free open source extension designed to perform ABM within the ArcGIS environment. One of the main goals for the integration of a GIS with an ABM is to provide a realistic landscape that can change through time on which the ABM runs. Agent Analyst is a mid-level integration of ArcGIS and RePast. ArcGIS provides spatial tools to prepare input data, manages the real landscape, displays the simulation results for each time step, and hundreds of spatial operations that can be used in the ABM. RePast provides the agent control and the scheduler for the ABM.

We created several ABM's to test the integration of Agent Analyst to ensure the user can perform tasks on large realistic datasets. One particular model was a cougar movement model. We worked with the USGS in Flagstaff, AZ to explore the interaction of cougars with their landscape and encounters with people. The model was performed on a large real landscape dataset (10 GB) and the rules were derived from scientists and data collected from collared cougars. We are aware of several other ABM models that have been implemented within this framework.

We found the software performed reasonably well. However the software is quite difficult to use especially for non-programmers. There is no easy way to create the rules for an agent other than through code. Looking through the future, we are interested to see if it is possible to develop a generic ontology that could be used to define agent rules. From this ontology a graphical user interface might be created which would greatly widen the audience for ABMs. Second, agents are still moved in x, y space independent of the GIS and then linked back to ArcGIS by feature data updates. Third, it is somewhat difficult to access the Java ArcObjects from within the ABM. This has limitations for future extensions of the framework and for creating more specific models.

Through creating and using Agent Analyst we discovered that GIS and ABM integration is more synergistic and bi-directional than we anticipated. The GIS provides input for creating agents (e.g., a shapefile), creates changing landscapes, and provides spatial modeling tools that are used each time step of the simulations. The ABM models and tracks the state changes and the decision making of the agents on this changing landscape and utilizes spatial functions (e.g., how far am I from something) in the decision making process for each time step.

Our work on Agent Analyst and dynamic modeling has revealed a number of important software issues:

- how to store and manage many outputs from a simulations and scenarios
- how time is explicitly handled
- synchronizing input time series data with different time intervals
- tools to analyze the simulation results
- metrics to compare and evaluate different scenarios

We are increasingly coming to the view that the distinction between dynamic modeling and ABM is fuzzy. Issues that promote the fuzziness include, if in the dynamic modeling framework the modeler can change the feature attributes like in the Schelling model is this ABM? With iteration and neighborhood notation in Spatial Analyst cellular automata models can be created; is this ABM? If Spatial Analyst changes from a passive mode (e.g., can only change your own value) to an active mode (e.g., allow the processor change other cell values) a modeler can then model the movement of oil from the perspective of the spill exchanging concentrations based on a series of neighborhood locations simultaneously and can change cell values beyond the immediate neighborhood in a CA model; is this ABM? If features and rasters are processed randomly instead of sequentially then situations in which each object changes the decisions of other objects can be addressed since the object making the first move is randomly determined. By processing a raster randomly the modeler can also perform dynamic allocations such as siting 50 bird nests in the most suitable locations assigning restrictions such as no three nests can be within 100 meters of one another; is this ABM? Changing states of features and rasters, moving agents based on the characteristics at each location on the landscape (grazing models), and cellular automata are relatively easy to do in a dynamic modeling environment. To address ABMs where each discrete object is aware of how to make complex decisions and changes the decision making process through mechanisms such as learning and memory poses a greater challenge for a general dynamic modeling environment.

Is ABM a formal modeling procedure or a state of mind derived from the object-oriented coding constructs? Castle and Crooks (2007) summarizes several studies defining the functionality of an ABM as containing autonomy, heterogeneity, active, pro-active / goal-directed, reactive / perceptive, bounded rationality, interactive / communicative, mobility, adaptation / learning. ESRI wants to understand what is the necessary functionality to meet these requirements. We are interested in understanding how to implement non-linear interactions, stochastic events, and how modeling objects make trade offs between multiple criteria in decision making.

By formally defining ABM the requirements for creating ABM models will become more obvious. Clearly there is great overlap between dynamic modeling and ABM. What we at ESRI are most interested in is the essence of the functionality that is being asked for by both dynamic models and ABM. If we can define the essence of the functionality, we would like to create an integrated dynamic modeling environment that can hopefully encompass ABM.

Castle, C.J.E., and A.T. Crooks, 2006. Principles and Concepts of Agent-Based Modelling for Developing Geospatial Solutions, *UCL Working Papers Series*, London, UK.

GIS and Agent-Based Modeling of Complex Spatial Systems

A Position Paper for the Workshop on Agent-Based Modeling of Complex Spatial Systems

Naicong Li
The Redlands Institute, University of Redlands
naicong_li@institute.redlands.edu

It has been long recognized that it is difficult to describe and predict the behavior of a complex dynamic system with an analytical approach. In recent years, simulation with agent-based modeling has become an attractive alternative for studying complex dynamic systems. A complex system may be analyzed through simulation with a set of agents and their environment. Agents act, interact with each other agents, and react to their changing environment according to a set of behavioral rules derived from an underlying theory for the processes and interactions within a particular system. Increasingly realistic models for the behavior of complex systems can emerge from a cycle of simulation, validation and refinement.

Recent development in this area also includes the effort of integrating agent-based modeling with GIS to simulate dynamic spatial systems. Through this integration, researchers are able to incorporate detailed real-world environmental data, to simulate agent behaviors and processes as change and movement conditioned by GIS data representations of space and geography, and to visualize the results in 2D or 3D GIS environment. Furthermore, agent-based models can include real-time GIS data feeds to simulate and visualize situations unfolding in real time.

The Redlands Institute at the University of Redlands has been conducting experiments with the integration of agent-based modeling and GIS technology. We are exploring the potential that this technology integration could offer to the research of modeling and simulating complex spatial systems, as well as to identifying and addressing the gaps that still exit in meeting such research needs. Technologies under consideration include ESRI's *ArcGIS*® coupled with *Repast* (Recursive Porous Agent Simulation Toolkit[1]) via *Agent Analyst* (an ArcGIS extension), and *Cinema 4D*. We developed several experimental prototype models for simulating urban growth, creature movement (such as predator and prey interaction), military maneuvers, bird migration, and other phenomena such as hurricanes, atmospheric particle dispersion, line of sight terrain reasoning, etc.

At the conceptual representation level, these prototype, spatially-enabled / agent-based models combine both object and field modeling. The environments in these models are typically simulated using raster layers from a GIS database, each representing a particular environmental factor, or data surfaces derived through geoprocessing. The agents are typically simulated through GIS vector data objects, such as points (e.g. for moving

[1] REPAST. http://repast.sourceforge.net/

creatures) and polygons (e.g. for parcels). While an 'agent' object is not yet a supported notion in ArcGIS itself, Agent Analyst, through its connection to Repast, provides a framework for easily defining various agent types, their properties, and their behavioral rules. Actual GIS data can easily be "cast" into these agents and their respective property values at a particular state.

To simulate dynamic processes, the agents in these models are able to change (at each simulation interval) their spatial location (thus affecting movement), their orientation and view point, their movement speed, their shape, and the state of their other internal properties. These changes are conditioned programmatically based on behavioral rules which 'sense' and 'adapt' dynamically to the current state of other relevant agents, and in response to the current state of the environment. The environments in these models may change dynamically based on the time factor and/or as a result of the agents' actions. The new state (location, size, various attribute values, etc.) of each of the agents is updated on the GIS data surface, and is visualized in ArcGIS dynamically. Depending on the rate of visualization, these models can simulate movement and state change in near-real time. In effect the role of GIS in this case is to visualize the new state of the model at each simulation interval, while the actual model with the agents and their current state at any given time are maintained in Repast. The actual execution of agent behavior rules is also done in Repast, but it is possible to pass some of the spatial analysis processing to ArcGIS. The state of each simulation interval could also be saved as time stamped data, producing a set of time series data that can be used after the simulation is finished, to be replayed in ArcGIS, or to be presented in dynamic graphs (based on a chosen parameter), or to be visualized in some other tool, such as ArcGIS Tracking Analyst (to visualize the movement trajectory dynamically).

To better represent and simulate various geographic processes, we have also tried to identify generic components that are common to all the processes of a similar type. We have done a preliminary analysis of agent movement. We have identified basic movement types, by defining their movement path specification in terms of some more primitive concepts including: location; distance; direction; the moving agent's orientation; topological relations between the movement trajectory and reference objects; the surrounding environment; speed; and temporal sequence of sub-movement processes within a complex movement type. The formalization of this study into an ontology is still work in progress.

Agent-based modeling of complex spatial systems is one of our main research areas at the Redlands Institute. I have personally participated in this research effort. I hope to have the opportunity to attend this workshop, meet other scientists of the relevant fields and learn about their work, and participate in the discussions on the current issues in this research area.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Infusing geodemographics into agent-based models of social systems

Paul Longley, University College London

I am something of an interloper at this meeting, in that although I co-supervise two Ph.D. students with core interests in agent based modelling (Christian Castle and Andrew Crooks), my research is focused more around understanding and modelling socioeconomic distributions. I am a co-PI with Mike Batty of the UCL GeoVUE (Geographic Virtual Environments) Node of the UK National Center for E Social Science, and have worked with Richard Milton on the display of various geodemographic representations using the Google Map Creator. Christian's work is focusing upon the development of a pedestrian evacuation model for London's King's Cross station (the new terminal for the international Eurostar service) using Repast. The work investigates a range of scenarios, and breaks new ground in the representation of the socioeconomic characteristics of users of the transit system at different times of the day and week. Some differentiation is achieved through representation of different mobility characteristics (e.g. week-end travellers with luggage versus weekday commuters) overlain with different demographic characteristics and variable familiarity with the internal structure and configuration of the underground structure.

Agent-Based Modeling Position Statement

Kevin M. Johnston and David Maguire (ESRI)

This position paper is written from the perspective of ESRI as a developer of commercial-off-the-shelf GIS products that ship in volume. Our overarching goal is to build software that solves important problems in a widely applicable way. As such, solutions must be robust, reliable, generic, deliver acceptable performance on large, real world data sets, and be easy to implement.

ESRI has 25 years of experience creating spatial modeling tools for vector and raster analysis. In the last five years we have built a geoprocessing modeling environment that synthesizes our experience in spatial modeling into an easy to use, flexible framework. Within the framework hundreds of spatial modeling functions are available each as individual tools that can be accessed from dialogs, the command line, a graphical modeling environment (ModelBuilder), scripting, or from any programming language. The framework is fully object–oriented and is built on .Net standards. In using ModelBuilder a model is comprised of data that are connected to spatial and aspatial tools to create a process to perform some function. Processes can be linked to one another to create simple as well as complex models. This framework has been used extensively to create static, cartographic process models that encompass classic map analysis/algebra operations.

In the latest 9.2 release of ArcGIS we expanded the geoprocessing environment, in particular ModelBuilder, with looping capabilities and the ability to model stochastic events through a series of random number and data creation functions. With these new capabilities we have explored how to add time to models and how to create various simulation scenarios in a dynamic modeling environment. We have been particularly interested in how to develop process models (e.g., a fire growth model), process models with stochastic events (e.g., a fire growth model with spotting), error analysis through simulations (e.g., examine the effects of error in a DEM on a stream delineation model), and sensitivity analysis (e.g., how does the output change with a 5% change in the distance to roads in a housing suitability model).

With a strong interest in agent-based modeling, we worked with Argonne National Laboratories, the creators of the RePast ABM system, to build Agent Analyst. Agent Analyst is a free open source extension designed to perform ABM within the ArcGIS environment. One of the main goals for the integration of a GIS with an ABM is to provide a realistic landscape that can change through time on which the ABM runs. Agent Analyst is a mid-level integration of ArcGIS and RePast. ArcGIS provides spatial tools to prepare input data, manages the real landscape, displays the simulation results for each time step, and hundreds of spatial operations that can be used in the ABM. RePast provides the agent control and the scheduler for the ABM.

We created several ABM's to test the integration of Agent Analyst to ensure the user can perform tasks on large realistic datasets. One particular model was a cougar movement model. We worked with the USGS in Flagstaff, AZ to explore the interaction of cougars with their landscape and encounters with people. The model was performed on a large real landscape dataset (10 GB) and the rules were derived from scientists and data collected from collared cougars. We are aware of several other ABM models that have been implemented within this framework.

We found the software performed reasonably well. However the software is quite difficult to use especially for non-programmers. There is no easy way to create the rules for an agent other than through code. Looking through the future, we are interested to see if it is possible to develop a generic ontology that could be used to define agent rules. From this ontology a graphical user interface might be created which would greatly widen the audience for ABMs. Second, agents are still moved in x, y space independent of the GIS and then linked back to ArcGIS by feature data updates. Third, it is somewhat difficult to access the Java ArcObjects from within the ABM. This has limitations for future extensions of the framework and for creating more specific models.

Through creating and using Agent Analyst we discovered that GIS and ABM integration is more synergistic and bi-directional than we anticipated. The GIS provides input for creating agents (e.g., a shapefile), creates changing landscapes, and provides spatial modeling tools that are used each time step of the simulations. The ABM models and tracks the state changes and the decision making of the agents on this changing landscape and utilizes spatial functions (e.g., how far am I from something) in the decision making process for each time step.

Our work on Agent Analyst and dynamic modeling has revealed a number of important software issues:

- how to store and manage many outputs from a simulations and scenarios
- how time is explicitly handled
- synchronizing input time series data with different time intervals
- tools to analyze the simulation results
- metrics to compare and evaluate different scenarios

We are increasingly coming to the view that the distinction between dynamic modeling and ABM is fuzzy. Issues that promote the fuzziness include, if in the dynamic modeling framework the modeler can change the feature attributes like in the Schelling model is this ABM? With iteration and neighborhood notation in Spatial Analyst cellular automata models can be created; is this ABM? If Spatial Analyst changes from a passive mode (e.g., can only change your own value) to an active mode (e.g., allow the processor change other cell values) a modeler can then model the movement of oil from the perspective of the spill exchanging concentrations based on a series of neighborhood locations simultaneously and can change cell values beyond the immediate neighborhood in a CA model; is this ABM? If features and rasters are processed randomly instead of sequentially then situations in which each object changes the decisions of other objects can be addressed since the object making the first move is randomly determined. By processing a raster randomly the modeler can also perform dynamic allocations such as siting 50 bird nests in the most suitable locations assigning restrictions such as no three nests can be within 100 meters of one another; is this ABM? Changing states of features and rasters, moving agents based on the characteristics at each location on the landscape (grazing models), and cellular automata are relatively easy to do in a dynamic modeling environment. To address ABMs where each discrete object is aware of how to make complex decisions and changes the decision making process through mechanisms such as learning and memory poses a greater challenge for a general dynamic modeling environment.

Is ABM a formal modeling procedure or a state of mind derived from the object-oriented coding constructs? Castle and Crooks (2007) summarizes several studies defining the functionality of an ABM as containing autonomy, heterogeneity, active, pro-active / goal-directed, reactive / perceptive, bounded rationality, interactive / communicative, mobility, adaptation / learning. ESRI wants to understand what is the necessary functionality to meet these requirements. We are interested in understanding how to implement non-linear interactions, stochastic events, and how modeling objects make trade offs between multiple criteria in decision making.

By formally defining ABM the requirements for creating ABM models will become more obvious. Clearly there is great overlap between dynamic modeling and ABM. What we at ESRI are most interested in is the essence of the functionality that is being asked for by both dynamic models and ABM. If we can define the essence of the functionality, we would like to create an integrated dynamic modeling environment that can hopefully encompass ABM.

Castle, C.J.E., and A.T. Crooks, 2006. Principles and Concepts of Agent-Based Modelling for Developing Geospatial Solutions, *UCL Working Papers Series*, London, UK.

Complications of complexity in agent-based models
George Malanson, University of Iowa

I have been working on both computational aspects of complex systems and agent-based models. My background is primarily in simulating vegetation dynamics using individual-based models (ibm) and cellular automata. This work includes analysis of the computational issues that arise when attempting to use evolutionary algorithms to simulate spatial systems that may exhibit self-organization and the difficulties of simulating agent decision making with too little data or not the right data to specify the decision. I am also involved in agent based modeling of land use decisions both directly and indirectly in ongoing funded research.

My perspective is that systems in which feedbacks occur between the spatial patterns and the drivers creating them are thus nonlinear, and varied spatial patterns can be produced by simple nonlinear processes. These "emergent" patterns that develop are at a scale (spatial, temporal or phenomenological) larger than the processes can be trivial unless they can be interpreted. Meaningful interpretation will depend on first better quantification of the dynamics of pattern and second, necessarily, on the development of an explanatory narrative. However, most systems of interest are both nonlinear and complicated. A real challenge is to determine to what degree the evolution of patterns is determined by the complications or the nonlinearity.

Models of agents are one way in which complexity and complication can be studied and perhaps differentiated, but they run the risk of confounding the two. Additionally, for the study of real places and people, models based on econometrics will be limited and the challenge is to bring narrative information into agent simulations. As well-known in ecological ibm, the initial locations of individuals strongly affects the outcome, and many agent models are not good at assigning calibrated agents to the right place (in some applications privacy issues will actually prevent this). The outcomes of agent models may then display results, especially "emergence," that are at scales that are too coarse and/or general to help understand real places – although they may allow comparison with other dynamical systems and identify system constraints, which could be useful.

Key insights that I can bring to the table include (references cited are on my bio pages):

In self-organizing systems, evolutionary computation has the potential to help understand the broad form of the functions describing system behavior, but self-organization limits or perhaps prohibits the determination of a narrow range of function specifications (Malanson & Zeng 2004).

Local, nonlinear pattern-process feedback can extend to broader scale linear correlations between pattern and process This type of interaction indicates self-organized complexity, but the form is yet to be determined (Zeng & Malanson 2006).

The introduction of exogenous drivers of pattern formation into a self-organzing system produce a threshold response where the size of the patterns exceeds the spatial extent of the self-organizing feedback (Zeng, Malanson, Butler, in review).

GIScience and landscape ecology can develop synergies by building on this area of geocomputation and complexity theory, as in analysis of attractors in state spaces of spatial metrics from spatially explicit simulations and representing their uncertainty; visualization is insufficient (Malanson et al. 2006a).

Power-law distributions and/ or alternative approaches in self-organized complexity, including self-organized percolation and the inverse cascade model, and highly optimized tolerance, based on their common ancestry in percolation theory, might provide insights into spatial pattern development (Malanson et al. 2006b).

High-resolution constraints on human actions and interactions in space and time

Harvey J. Miller
Department of Geography
University of Utah

A recent paper by Boman and Holm (2004) argues for the convergence of three approaches for modeling human spatial systems from the "bottom-up": i) microsimulation; ii) agent-based modeling (ABM), and; iii) time geography. These approaches offer complementary strengths for analyzing and understanding human phenomena. Microsimulation offers a tradition of computational approaches to understanding how micro-level behavior creates dynamic human phenomena, as well as standards for model estimation and validation. However, microsimulation models typically represent human behavior in an aggregate and isolated manner since behaviors manifest from cohorts rather than individual actions, interactions among humans, as well as interactions between humans and the environment. In time geography, interactions among humans and with the environment are fundamental, but linkages between individual behaviors and aggregate social and environmental dynamics are only conceptual in nature. ABM offers a rigorous but rich approach to simulating human phenomena from the bottom-up, as well as the concepts of adaptation, self-organization and emergence to capture linkages between individual behavior and aggregate dynamics. ABM can benefit from time geography's focus on constraints, as well as microsimulation's adherence to estimation and validation standards.

While time geography has much offer as a conceptual foundation for ABM, a weakness is its traditional lack of a rigorous analytical foundation. Time geographic entities such as the path and prism, and relationships such as path bundling, path-prism intersections and prism-prism intersections, are described only informally. Classical time geography is not sufficiently developed to support measurement and analysis using high resolution technologies such as ABM, as well as data collection using *location-aware technologies* such as the global positioning systems (GPS) or radiofrequency identification (RFID) chips.

Rigorous, high resolution constraints on human activities and interactions in space and time are possible through temporal disaggregation of time geography. At a given moment in time, entities such as the space-time path and prism are simple geometric objects in the low-dimensional space of interest in time geography. For example, the space-time path at a given moment in time is a point object. Further, the space-time prism at a given moment in time is the intersection of three compact spatial sets that have simple geometric forms. Given the low dimensional space of interest in time geography, it is easy to solve for the prism, prism-prism and path-prism intersections, as well as other relationships such as path bundling (Miller 2005a). In addition to allowing easy solutions, the temporal disaggregation meshes well with the discrete temporal dynamics in microsimulation, ABM and LATs.

The high-resolution time geographic measurement theory can also be extended to encompass virtual interaction. Using Janelle's typology of interaction constraints based on spatial presence versus telepresence and temporal synchronicity versus asynchronicity as a foundation (Janelle 2004), we introduce new time geographic objects: i) a *portal* (locations that allow virtual interaction), and ii) *message windows* (potential or actual communication events represented as time intervals). Using the well-known Allen time predicates that encompass all possible relationships between two intervals of time, it is possible to derive the rigorous constraints on communication events within the Janelle framework (Miller 2005b).

More recently, we have extended the time geographic measurement theory to the case where travel velocities vary continuously across space. Using the continuous transportation modeling or "urban fields" theoretical tradition in geography and regional science, we formulate analytical definitions of the space-time path and prism for the case where unobserved components are characterized by minimum cost curves through an inverse velocity field rather than straight line segments based on a uniform velocity. The theory is more realistic and also links time geography to the continuous space modeling tradition in geography and regional science. In addition to theoretical relevance and elegance, the time geographic field approach provides a synoptic summary and visualization technique that can provide insights to space-time accessibility patterns. Preliminary results suggest that the space-time prism is sensitive to the velocity assumption, and the traditional prism based on a uniform maximum velocity assumption is only a special case of a family of prisms with a wide range of geometric forms.

Literature cited

Boman, M. and Holm, E. (2004) "Multi-agent systems, time geography and microsimulations," in Olsson, M.-O. and Sjöstedt, G. (eds.) *Systems Approaches and their Applications* Dordrecht: Kluwer Academic, 95-118.

Janelle, D. G. (2004) "Impact of information technologies," in S. Hanson and G. Giuliano (eds.) *The Geography of Urban Transportation,* 3ed., New York: Guilford, 86-112

Milller, H. J. (2005a) "A measurement theory for time geography," *Geographical Analysis*, 37, 17-45

Miller, H. J. (2005b) "Necessary space-time conditions for human interaction," *Environment and Planning B: Planning and Design*, 32, 381-401.

NSF/ESRC Agenda Setting Workshop on <u>Agent-Based Modeling of Complex Spatial Systems</u>: April 14-16, 2007

Richard Milton
CASA, UCL

Computational Representations and Dissemination of Fine Scale Geographic Data

I am working on representing geographic data associated with spatial analysis and modelling in forms that enable a wide variety of users to access analytical information. This information is in a form that can be mapped thematically as well as information on movement that can be animated. Although the technologies we are working on are independent of scale (within obvious limits), the recent focus of this work is on fine scale data required for agent-based modelling such as remotely sensed data from GPS in various contexts ranging from monitoring pollution to tracking children moving from home to school.

We are building various tools to display such data in non-proprietary web-based systems such as *Google Maps* and *Google Earth*. In this note, I will explain our work with *Google Maps* which involved the development of a freeware application to aid with the production and publishing of maps on the web. If the data involves movement on the map we can display this in a form that links the movement to other critical variables such as pollution and energy levels. This project is part of our work on developing new computational resources for analysis and modelling in terms of display and dissemination which are currently web based services with the potential to utilise grid based services when the size of the problem exceeds certain thresholds.

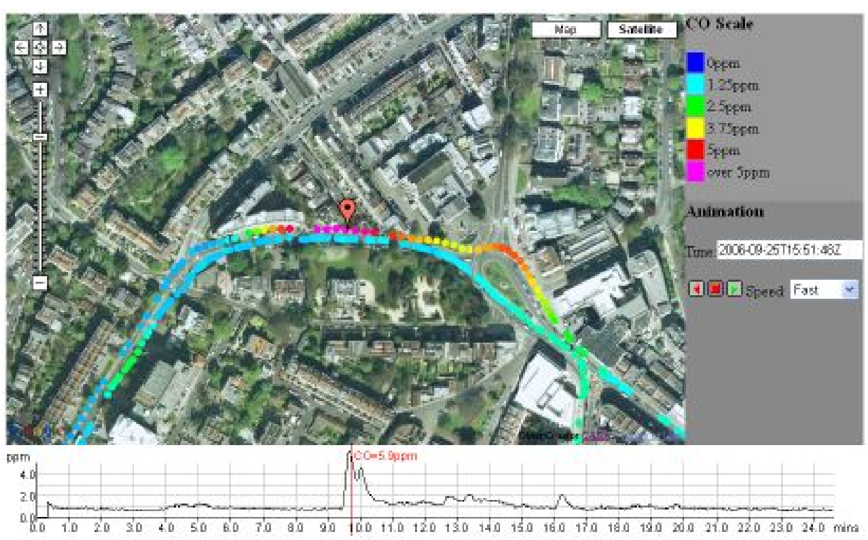
- The Google Map Creator is a Java application designed to make publishing thematic data on Google Maps easier. The application takes data in the form of a shapefile and colours the areas according to a user defined colour scale. All the Google Maps tiles for the area covered by the data are rendered and saved to disk, along with an html file. The result is a working web site that can be copied to a web server for publishing on the internet. The key feature of this software is that it allows maps to be published by a wider range of users than was previously possible. The result is a completely file based site that does not rely on a web service to create the tiles, so maps can be published by anyone with access to web space that they can copy files to.

  The software works with shapefiles containing data in any projection, as long as there is a valid projection file. This is achieved by using the Geotools library to do the necessary reprojection into the two projections needed to create the tiles for the Google Map. Once a colour scale has been chosen by the user, the features are rendered to all the tiles covered by the data. This is done outside of Geotools, but using some of the Geotools functions to access and transform the data, along with some of the spatial indexing functionality built into JTS. The time taken to render all the tiles can be very long, so this part of the software had to be threaded to enable the user interface to work while the tile rendering is taking place.

- Another application that we developed is the Google Maps Image Cutter. This application takes an image and displays it on the web using Google Maps. Using Google Maps in this way allows the publishing of very large images such as panoramas and gives the user the ability to pan over the image and zoom in to see more detail.

- Work has also been done on animating GPS tracked data and publishing it on the web. This mainly involves GPS tracked carbonmonoxide data from a previous project, but we were also involved in the filming of the BBC programme 'Don't Die Young' (screened on 23rd January 2007) where a cyclist was tracked through Bristol with a carbonmonoxide sensor. The data is published on the CASA web site at the following address:

  http://www.casa.ucl.ac.uk/bbc/dontdieyoung/log_25-09-2006_154206.html



Further Information about Google Map Creator and related software systems is available from:

http://www.casa.ucl.ac.uk/software/googlemapcreator.asp

http://www.casa.ucl.ac.uk/software/googlemapimagecutter.asp

Reference:
Milton, R., and Steed, A. (2007) Mapping Carbon Monoxide Using GPS Tracked Sensors, *Environmental Monitoring and Assessment*,
DOI 10.1007/s10661-006-9488-y

Position paper for UCSB workshop "Agent-based modeling of complex spatial systems"
Dawn Parker, George Mason University
Feb. 19, 2007

Early successful examples of spatial agent-based models were often highly abstract, and were used to demonstrate how particular macro-scale outcomes could emerge as the result of decentralized interactions of autonomous decision-making agents. Examples include the demonstration of segregation provided by Schelling's early models (Schelling 1971) and the emergence of power-law distributions of wealth from Epstein and Axtell's Sugarscape model (Epstein and Axtell 1996). Such models had a very simple role: to extend theory by demonstrating that a particular set of behaviors and interactions could *generate* a particular macro-scale or emergent outcome (Epstein 1999). As the discipline of spatial-agent based modeling matures, an increasing number of highly empirical models (in the style of "Cell 4" models as described by Parker, Berger, and Manson (2002) are being constructed. These models are often developed to support policy analysis, and are motivated by a belief that models that include complex dynamics and interactions are likely to produce "better" results that models that omit those dynamics. While my discussion focuses on models of human-environment interactions (Bousquet and Le Page 2004; Grimm and Railsback 2006; Janssen 2003; Parker et al. 2003), such empirical spatial agent-based models are also being developed in other areas of social science, including epidemiology, sociology, and political science (Castle and Crooks 2006).

Many scholars have voiced concern that such highly empirical models of complex systems are in danger of being black boxes in which the relationships between inputs and outputs, and even the behavior of the model itself, are not well understood. Some argue that for these reasons agent-based modeling is appropriate only for highly abstract demonstration of theoretical outcomes, and that the push to increase complexity in these empirical models may run counter to scientific principles of parsimony in modeling. Fundamental questions are raised about where and how agent-based modeling fits into the scientific method. In the contrary view, others argue that the policy environment is complex and therefore demands development of complex models, and that policy questions we now face, such as potential contributions of land-use and land cover change to global climate change, and the potential transmission paths of pandemics, are so pressing that any and all practical approaches to obtaining answers must be tried. A separate thread argues that spatial agent-based models may contribute to the development of an interdisciplinary theory of spatial social science. Yet, whatever one's view of the appropriate role for ABM in spatial social science, tensions exist between the need to simply build better empirical models (that better represent real-world drivers and complex interactions) and goal of building theoretical frameworks that have a more formal relationship to the scientific method, which could then contribute to building integrated theories of spatial social science. The first goal demands more details and realism; the second demands a high-level abstract framework that is generalizable across case studies and perhaps even realms of social science.

The questions lain out in this position paper attempt to steer the debate from what sorts of spatial agent-based models we should be doing to how we might do modeling—both theoretical and empirical. These questions also aim to focus on how agent-based modeling can be more closely connected to the scientific method, at both theoretical and empirical ends of the spectrum. While empirical models face additional challenges of parameterization, calibration and validation, models constructed at either end of the spectrum share a goal of incorporating as much complexity is needed to represent the problem under study, but not too much. And both types of models ultimately seek to make substantive contributions to science.

Question 1: How to identify a just-sufficient level of detail for the model? The idea that the goal of agent-based models is to produce macro-scale emergent patterns through micro-level behaviors and interactions is well established. This idea can, however, provide insufficient guidance to create a parsimonious model. A slight refinement of the the model question to "How do particular agent behaviors and interactions at a micro-level produce patterns observed at a macro-level?" can help focus efforts. Grimm et al. (2005) suggest that at a minimum, your model must embed processes of sufficient complexity such that it produces the macro-scale patterns of interest. Further, these patterns must of course be non-trivially produced through micro-scale interactions, rather than be obvious consequences of the rules specified at an agent level. Initially, of course, a modeler will introduce micro-level behaviors and interactions that she believes are linked to macro-scale patterns. In relation to land-use modeling, I characterize these behaviors and interactions into spatial, temporal, and behavioral complex drivers (Parker 2007). Grimm et al. (2005) also note that because multiple processes can produce the same single observed pattern, often multiple observed output patterns are needed to distinguish between competing process models. In short, identifying a minimum amount of complexity requires a clear initial hypothesis that links a micro-scale exogenous driver (which could be a model parameter, an agent behavior, or the structure of agent interactions) to a macro-scale emergent pattern outcome. A focus on linkages between hypothetical drivers and outcomes, rather than simply on reproduction of observed patters, could reunify the process of model building with the traditional scientific method. Further, for highly empirical models designed for policy analysis, it brings a focus to policy levers whose values may be modified for scenario or sensitivity analysis, thus encouraging the development of the model to be driven by its intended use.

Question 2: What are potential roles for statistical analysis in spatial agent-based models? In Parker et al. (2003), statistical models were characterized very much as substitutes for agent-based models. I have modified that view substantially since that paper was first written. I now see statistical methods as an essential part of the agent-based modeling process; certainly for highly empirical models, but also for purely theoretical models. In both cases, however, there are yet (to my knowledge) too few statistical tools available specifically for analysis of output from complex spatial systems. I strongly suspect that these tools exist and await only discovery and better communication between currently disconnected groups of researchers.

In addition to the roles for statistical models in building empirical agent decision functions described in Robinson et al (Forthcoming), I see two potential additional roles:

Application 1: Pseudo-inductive analysis. Demonstrating that an ABM can recreate an observed pattern is a first step towards using ABM as a substitute for traditional abstract mathematical models. Yet a modeler is likely to have larger goals. He may want to demonstrate that the outcome holds globally, over a large (and reasonable) range of parameter values. He may also want to understand global relationships between directions of change of parameters and directions of change of outputs. The solution to this problem has been well articulated in the agent-based modeling literature (Axelrod and Tesfatsion 2006): the modeler should create a database of outcomes by sweeping the parameter space, then use statistical methods to analyze the generated data. However, what has been less well articulated (to my knowledge) is what tools are available that are statistically appropriate for analysis of complex systems. By their nature, complex systems are characterized by endogeneity between micro-scale elements and macro-scale outcomes due to cross-scale feedbacks. They are also characterized by non-linear response surfaces and thresholds where abrupt changes occur. These properties violate the assumptions of the mostly-linear regression models in the historical toolkit of social science modelers. Although traditional regression techniques have been applied to the analysis of output from agent-based land-use models by myself (Parker 2005) and Happe et al (2006), both authors acknowledge the limitations of the linear models that they employ.

Application 2: Model validation. The first application focuses on statistical analysis of the relationship between micro-scale drivers and macro-scale measures of pattern outcome. A second potential application (not applied to the author's knowledge) involves statistical analysis of ABM output at the micro level, and comparison of that model output to results from the same statistical model applied to real-world data (which of course would have to lie outside any data used to build the ABM). For example, an empirically-parameterized agent-based model of residential land markets would produce outputs at the micro level similar to those used to estimate spatial econometric models of land-use change, including land-use transitions, transaction prices, distance to transportation networks, and neighborhood composition at multiple spatial scales. Such data could be used to estimate a land-use change model using techniques similar to those described by Bell and Irwin (2002). If the estimated parameters of the simulated and real-world models had the same signs and statistical significance, the model could be said to exhibit qualitative agreement with the real world-data. If the parameter estimates were statistically similar, the model would have hit the proverbial "home run" (although likely other validation methods would probably also be called for). This approach goes beyond asking if the model replicates spatial outputs such as location, pattern, and composition at multiple scales, and rather asks how closely the model replicates the structural relationships found in the real-world data. (Additional thought would be required to account for path-dependence and the distribution of simulated model outcomes that would be possible (Brown et al. 2005).)

Question 3: What is the role for calibration in empirical agent-based models?
Calibration (derivation of a set of best-fit model parameters through comparison of outcomes at the micro or macro level) has played multiple roles in spatial empirical ABMs. Both statistical and mathematical programming-based agent decision models have been calibrated using micro-level outcome data (Balmann et al. 2003; Berger 2001; Happe, Kellermann, and Balmann 2006; Schreinemachers and Berger 2006). Parameters of agent decision models have also been calibrated using macro-scale data on land-use composition and pattern (Caruso, Rounsevell, and Cojocaru 2005; Evans and Kelley 2004). Each of these methods implicitly assumes that the chosen structure of the decision function is correct, but that the parameters of that function are uncertain. Thus calibration does not necessarily provide a means to distinguish between competing decision models. If macro-scale pattern outcomes of interest are used for calibration, then separate outcome data must be used to validate the model after calibration, increasing overall data requirements for the model. Decision models often contain many more parameters than the number of observations available on macro-scale outcomes, leading to potential parameter identification problems, and exacerbating the problem of distinguishing between alternative decision models, as there may be cases where different sets of parameter values for the same decision model may lead to equivalent pattern outcomes. Grimm et al (2005), referring to the calibration process described above as "inverse modeling", propose calibration via multiple pattern outcomes as a means of addressing the parameter identification problem. Calibration may have an important role to play for spatial ABM at a later stage of model development and/or in cases where agent rules and behaviors are well known (for example, when they have been determined via field observation or laboratory experiments).

References:

Axelrod, R., and L. Tesfatsion. 2006. A Guide for Newcomers to Agent-Based Modeling in the Social Sciences in L. Tesfatsion and K. Judd, eds. *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*. North-Holland/Elsevier, Amsterdam

Balmann, A., K. Happe, K. Kellermann, and A. Kleingarn. 2003. Adjustment costs of agri-environmental policy switchings: A multi-agent approach in M. A. Janssen, ed. *Complexity and Ecosystem Management: The Theory and Practice of Multi-agent Approaches*. Edward Elgar Publishers, Cheltenham, U.K.; Northampton, MA

Bell, K. P., and E. G. Irwin. 2002. Spatially explicit micro-level modelling of land use change at the rural-urban interface. *Agricultural Economics* 27 (3): 217-232

Berger, T. 2001. Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes, and policy analysis. *Agricultural Economics* 25 (2-3): 245-260

Bousquet, F., and C. Le Page. 2004. Multi-agent simulations and ecosystem management: a review. *Ecological Modelling* 76 (3-4): 313-332

Brown, D. G., S. E. Page, R. Riolo, M. Zellner, and R. W. 2005. Path dependence and the validation of agent-based spatial models of land use. *International Journal of*

Geographic Information Systems 19 (2): 153-174. http://www.pscs.umich.edu/research/projects/sluce/publications/ijgis-sluce-final.pdf.

Caruso, G., M. Rounsevell, and G. Cojocaru. 2005. Exploring a spatio-dynamic neighbourhood-based model of residential behaviour in the Brussels periurban area. *International Journal of Geographical Information Science* 19 (2): 103-123

Castle, C., and A. Crooks. 2006. Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. London: CASA Publication 110. http://www.casa.ucl.ac.uk/publications/workingPaperDetail.asp?ID=110.

Epstein, J. M. 1999. Agent-based models and generative social science. *Complexity* 4 (5): 41-60

Epstein, J. M., and R. Axtell. 1996. *Growing Artificial Societies: Social Science from the Ground Up*. Brookings Institution Press, Washington, D.C.

Evans, T. P., and H. Kelley. 2004. Multi-scale analysis of a household level agent-based model of landcover change. *Journal of Environmental Management* 72: 57–72

Grimm, V., and S. F. Railsback. 2006. Chapter 1: Introduction in V. Grimm and S. F. Railsback, eds. *Individual-based Modeling and Ecology*. Princeton University Press, Princeton, NJ

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis. 2005. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* 310: 987-991

Happe, K., K. Kellermann, and A. Balmann. 2006. Agent-based Analysis of Agricultural Policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its Adaptation and Behavior. *Ecology and Society* 11 (1): 49. http://www.ecologyandsociety.org/vol11/iss1/art49.

Janssen, M. A., ed. 2003. *Complexity and Ecosystem Management: The Theory and Practice of Multi-Agent Approaches*. Edward Elgar Publishers, Cheltenham, U.K.; Northampton, MA

Parker, D. 2007. Class Web Site: Spatial agent-based models of human-environment interactions. http://mason.gmu.edu/~dparker3/spat_abm/spat_abm.html.

Parker, D. C. 2005. Agent-Based Modeling to Explore Linkages Between Preferences for Open Space, Fragmentation at the Urban-Rural Fringe, and Economic Welfare. Paper presented in the The role of open space and green amenities in the residential move from cities, Dec. 14-16, Dijon, France.

Parker, D. C., T. Berger, and S. M. Manson. 2002. Meeting the Challenge of Complexity: Proceedings of the Special Workshop on Agent-Based Models of Land-Use/Land-Cover Change. Santa Barbara: CIPEC/CSISS Publication CCR-3. http://www.csiss.org/maslucc/ABM-LUCC.htm.

Parker, D. C., S. M. Manson, M. A. Janssen, M. Hoffmann, and P. Deadman. 2003. Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. *Annals of the Association of American Geographers* 93 (2). http://www.csiss.org/events/other/agent-based/papers/maslucc_overview.pdf.

Robinson, D. T., D. G. Brown, D. C. Parker, P. Schreinemachers, M. A. Janssen, M. Huigen, H. Wittmer, N. Gotts, P. Promburom, E. Irwin, T. Berger, F. Gatzweiler, and C. Barnaud. Forthcoming. Comparison of empirical methods for building agent-based models in land use science. *Journal of Land-Use Science*

Schelling, T. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143-186

Schreinemachers, P., and T. Berger. 2006. Land use decisions in developing countries and their representation in multi-agent systems. *Journal of Land Use Science* 1 (1): 29-44. http://journalsonline.tandf.co.uk/openurl.asp?genre=article&issn=1747-423X&volume=1&issue=1&spage=29.

AGENT-BASED MODELING OF COMPLEX SPATIAL SYSTEMS WORKSHOP POSITION PAPER

Donna J. Peuquet
Penn State University

Virtually all current scientific, socioeconomic and security questions depend on geospatial information and on the ability of scientists to interact with that information in increasingly flexible and holistic ways, whether the problem context concerns, for example, managing urban growth, predicting the spread of disease, understanding an evolving world economy in the digital age or protecting against terrorist attack.

The important issues relating to these applications center on the understanding of geospatial *processes*, as entities change over time. The interactions among entities and their components within processes tend to be highly dynamic, interlinked, and with complex chains of cause-and-effect. For example, environmental water quality is affected by urbanization (housing and road density, etc.), which is in turn affected by public policy and economics governing how and where things may be built at both local and regional scales.

Because of the complex and non-deterministic interactions of social and natural components, improving our understanding of geospatial processes or deriving specific problem solutions generally requires representation and analysis of higher-level, domain-specific human knowledge in addition to observational data. The complexity of the processes being investigated also requires the application of expertise from multiple knowledge domains. How do the components of a process and the subcomponents of each relate to each other, and at varying spatio-temporal scales? Such questions are addressed via teams of analysts with differing domains of expertise and differing views of the phenomenon. Because of the limitations of current tools, their analyses tend to be unlinked and potentially in disagreement.

Approaches for representation and analysis remain focused on handling observational data. The current manifestation of this focus is the emergence of networks for the purpose of linking various data sources and software tools, and providing real-time access to these, via what has become known as the *cyber-infrastructure*. Higher-level, derived knowledge of geographic phenomena is rarely stored or utilized by the software. As pointed out by [1], GIScience has been focused on the study of *form* – of how world looks rather than how the world works. What can be derived from current geospatial software tools is thus limited by the expertise, experience, and memories of the specific people using the system at a given time. This can be problematic, particularly in an emergency situation when a solution must be quickly derived.

Formalized ontologies are widely seen as the means to overcoming this situation, and have received much attention within GIScience. Ontologies, in a digital context, provide a basis for augmenting data with a common semantics for database integration and sharing, as well as a stored knowledge base. These formalized ontologies are often described conceptually (and represented graphically) as concept graphs. While this approach provides a means of representing a level of abstraction above the observational data that is also intuitive for the user, it is not without serious shortcomings with regard to representing space-time process, as well as refining our knowledge of processes through new observational data.

Formalized ontologies tend to be built through a manual, custom-tailoring process. They may grow as users add concepts and relationships but the basic concepts and interrelationships among them are assumed to be static. In the real world, however, concepts evolve - the basic categorical groupings from which they derive change, concepts can be replaced or completely redefined, as well as the relationships among them – as knowledge grows.

OWL has quickly become the standard language for encoding formalized ontologies. OWL is a Web Ontology Language intended as a tool for providing a means for common understanding across the World Wide Web in the idea of the 'semantic net.' OWL was designed to define classes and concepts as well as their properties and relationships, and also to allow reasoning about these classes and concepts [2]. It, however, is weak in its operational ability to determine the need for change in the ontology represented via evidence from new data and to appropriately modify it. The volume of new geospatial data streaming in from various sources, and their heterogeneity (with the consequent complexity and variability of possible patterns) makes a focus on inductive techniques such as those associated with Data Mining insufficient. Theory-driven methods of analysis that are guided by stored knowledge, as well as observational data, are required.

As already stated, ontologies provide a specification of classes, concepts and the structure among them. Recent research has shown how this can be extended to include ontological history – explicit storage of new and superseded concepts, these linkages, and description of how components originate [3]. Nevertheless, the fundamental mechanisms of the underlying *process* – i.e., how things change and evolve to affect the form of the observed phenomenon and the derived, abstracted elements (or concepts) within that process – is not represented within any ontology. A process can be defined functionally as a function that acts on a domain and may or may not refer to, or act upon, observed objects applicable to the domain (a specific instance of, for example, a city, disease, or multinational corporation) stored within the data. The process could be simple and elemental (e.g., move forward), or could be more complex (shoreline erosion), in which case, it is composed from a dynamic set of hierarchically organized functions. Generally speaking, such a function incorporates a description of behaviors and their applicable context along with defined triggers to specific behaviors. Such a description would often consist of rules. At an elemental level, a function could be an algorithmic or mathematical transformation.

So, just as it is now acknowledged that storing large amounts of heterogeneous geospatial data requires a multi-representation framework, storing higher-level information concerning geospatial phenomena and their dynamics also requires a multi-representation approach. Moreover, these multiple data and knowledge representations must be functionally interlinked. Initial research at Penn State [4] has shown that the use of multiple software agent types linked with an ontology and database utilizing a what/when/where schema provides the needed combination of representational power.

There has been growing recent interest in intelligent software agents in GIScience as a specific tool in a variety of contexts, including simulation of land-use/cover change [5], wayfinding [6], and social simulations [7]. Software agents are not only a natural way to represent social aspects (institutions, societies, etc.) as well as natural aspects (climate, hydrology, etc.) of geospatial processes, but also to represent the coordinated dynamic behaviors of multiple entities constituting complex dynamic processes. They, in combination, can be applied in simulations of space-time processes to provide a key means of representing a specific process (or components of a process at a specific scale) in a dynamic way that deals with complex interactions [8].

Cognitive agents, as opposed to reactive agents, maintain an internal state (e.g., goals, plans, and state information about the world), and behave by searching through a space of behaviors and cooperating with other agents (which in turn may affect their actions)[9]. Key properties of cognitive agents with these characteristics are that besides being *situated* (applicable, or active, within specific contexts) and *distributed* (maintain their own constrained view), they are *adaptive* (learn and improve through experience). Cognitive agents can thus also be viewed as corresponding to the expertise of a human expert. Reactive agents tend to be more algorithmic in nature and are thus appropriate for dealing with low-level functions. A middle form (not fully cognitive) can also be useful as a flexible means of representing, various components of those processes.

Such a multiple agent scheme provides a means of representing knowledge about process (captured through simulation and verified via linkage with observational data, or explicitly input by the user). Agents, in combination with ontologies, act as knowledge carriers to provide qualitative information, derivation of higher-level abstractions, and makes large stores of observational data more easily retrievable, reusable, and shareable. This provides for a powerful simulation environment for studying space-time dynamics, but more importantly, such simulations provide a means of analyzing "what-if" alternatives to aide the human analyst/decision maker, and a fast means of conveying domain-specific knowledge to decision-makers in time-critical situations.

References

1. Goodchild, M., *GIScience, Geography, Form, and Process.* Annals of the Association of American Geographers, 2004. **94**(4): p. 709-714.
2. Aref, M.M. and Z. Zhou. *The Ontology Web Language (OWL) for a Multi-Agent Understating System*. in *International Conference on Integration of Knowledge Intensive Multi-Agent Systems (KIMAS 2005)*. 2005. Boston: IEEE.
3. Brodaric, B., *Representing Geo-Pragmatics*, in *Ph.D. thesis, Department of Geography*. 2005, Penn State University: University Park, PA.
4. Yu, C., *Geoagent-based Knowledge Systems, Ph.D. thesis, Department of Geography*. 2005, Penn State University: University Park, PA.
5. Parker, D.C., T. Berger, and S.M. Manson, *Agent-Based Models of Land-Use / Land-Cover Change*, in *LUCC Report Series No.6*. 2001, LUCC Focus 1 Office Anthropological Center for Training and Research on Global Environmental Change, Indiana University.
6. Raubal, M., *Ontology and epistemology for agent-based wayfinding simulation.* International Journal of Gegraphical Information Science, 2001. **15**(7): p. 653-665.
7. Conte, R., et al., *Sociology and social theory in agent based social simulation: a symposium.* Computational and Mathematical Organization Theory, 2001. **7**: p. 183-205.
8. Jennings, N.R., *On Agent-Based Software Engineering.* Artificial Intelligence, 2000. **117**(2): p. 277-296.
9. Ferber, J., *Multi-agent systems: an introduction to distributed artificial intelligence*. 1999, New York: Addison-Wesley.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Why Agent-Based Modelling of Complex Spatial Systems Needs Cyberinfrastructure.

Edoardo Pignotti
University of Aberdeen
Department of Computing Science
Aberdeen AB24 3UE, Scotland, UK
epignott@csd.abdn.ac.uk

Collaborations between large groups of scientists are increasingly seen as essential to enhance the scientific process. While research has always involved collaboration between individual scientists, there is now even greater necessity for tools to support sharing of knowledge, resources, results and observations.

The vision of e-science[1] (or Cyberinfrastructure) is to facilitate large scale science using Grid technologies as a fundamental computing infrastructure to manage distributed computational resources and data. However, a major gap exists between current technologies and the vision of e-science. Where Grid technologies overcome some of the limitations of existing Web tools in terms of managing computational tasks, there is still a need for greater ease of use and seamless automation to support truly flexible collaboration. For these reasons the concept of a Semantic Grid[2] has emerged, which integrates Semantic Web[3] and Grid technologies.

Central to the vision of the Semantic Grid is the adoption of metadata and ontologies to describe resources, services and data sources in order to promote enhanced forms of collaboration among the scientific community. Ontologies and metadata facilitate intelligent search mechanisms, one of the key enablers through which such services could be realised. The Semantic Web is a vision in which today's Web will be extended with machine readable content, and where every resource will be marked-up using machine readable meta-data. Ontologies are used to capture the meaning of meta-data terms and their interrelationships. The main benefit of using ontologies is that they facilitate access to heterogeneous and distributed information sources by defining a machine-processable semantics for those information sources. We argue that the Semantic Web approach has significant potential within eSocial Science. Semantic Web technologies can help deliver the vision of a more "human-centred" Grid which facilitates tasks such as collaboration, shared experimentation, and annotation of resources. Furthermore, we argue that these technologies have particular strength in capturing qualitative scientific arguments, supported by a mix of quantitative and qualitative data and results.

Recent activities in the field of social simulation have outlined the need to improve the scientific rigour of agent-based modelling. One of the important characteristics of scientific research is that work should be repeatable and verifiable. Yet results gathered from possibly hundreds of thousands of simulation runs cannot be reproduced conveniently in a journal publication. Equally, the source code of the simulation model, and full details of the model parameters used are also not journal publication material. We have identified several activities that are relevant:

- access to the results data itself, to check that the authors' claims that are based on those results are justifiable.
- an ability to re-run experiments to check that they produce broadly the same set of results.
- manipulation of the simulation model parameters to check that there is no undue sensitivity of the results to certain parameter settings.
- modification of the source code and/or re-implementation of the model to check for what might be termed 'algorithmic sensitivity'.

In a previous project, FEARLUS-G[4], we tried to bring together the needs of agent based modelling with the vision of eScience. This project involved social scientists at the Macaulay Institute in Aberdeen investigating

[1] http://www.rcuk.ac.uk/escience/default.htm
[2] http://www.semanticgrid.org/
[3] http://www.w3.org/2001/sw/
[4] http://www.csd.abdn.ac.uk/research/fearg/

land-use change and computer scientists at the University of Aberdeen. FEARLUS-G aimed to provide scientists interested in land-use phenomena with a means to run much larger-scale experiments than previously possible on standalone PCs, together with a Web-based environment in which to share simulation results. A key facet of this project was the development of an ontology which describes the tasks and entities involved in simulation work, such as experiments, hypotheses, parameters, simulation runs, and statistical procedures.

In the PolicyGrid[5] Project (a UK eScience project funded as part of the Economic & Social Research Council eSocial Science initiative) we are building on the work of FEARLUS-G. The project involves collaboration between computer scientists and social scientists at the University of Aberdeen, the Macaulay Institute (Aberdeen) and elsewhere in the UK. The project aims to support policy-related research activities within social science by developing appropriate Grid middleware tools which meet the requirements of social science practitioners. The project is developing a range of services to support social scientists with mixed-method data analysis (involving both qualitative and quantitative data sources) together with the use of social simulation techniques. Issues surrounding usability of Semantic Grid tools are also a key feature of PolicyGrid, with activities encompassing workflow support and natural language presentation of metadata.

As part as PolicyGrid we are investigating the use of semantic workflow tools to facilitate the design, execution, analysis and interpretation of simulation experiments and exploratory studies, while generating appropriate metadata automatically. We have explored a number of case studies of social simulation research activities. Some of the emerging challenges are:

- Capturing the scientist's goals and constraints associated with a workflow;
- Integrating workflows into a scientific argument structure;
- Improving interoperability and re-use of workflows.

To date, we have created an initial social simulation classification ontology capturing the characteristics of a wide range of simulation models, e.g. type of simulation, behaviour, space model, execution type, etc. Collaborators at the Macaulay Institute are continuing work on the development of a simulation modelling ontology to allow a particular piece of modelling software to be described and the structure and context of a particular simulation run to be characterised.

Questions

- Should there be a more systematic way of sharing and preserving the knowledge produced by the agent-based modelling community?

- Can we increase the scientific value of agent-based models by making them easier to verify and validate?

- Can workflow tools play an important role in capturing agent-based modelling methodology?

- What kind of activities are undertaken in agent-based modelling?
    - What are the relationships between these activities?
    - What kinds of metadata support do they need?

- Is it possible to draw up a general specification of metadata requirements for agent-based modelling, or do agent-based modelling projects differ in too many respects for this to be feasible?

- What software tools are needed (or would be useful) to support the activities listed, and the generation of the required metadata relevant to each activity?

[5] http://www.policygrid.org

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Victor Roger Schinazi
CASA, UCL

The idiosyncrasy of movement: GIS as a tool for exploring individual differences

The concept of individual differences continues to gain considerable attention within the discipline of psychology. A recent workshop at the 2006 *Spatial Cognition* conference (in Bremen, Germany) made evident that much can be learned by focusing on the individual and the particularities of both planned and unplanned navigation. Individuals vary in their ability to learn and navigate large-scale environments. These differences however are often treated as secondary, aggregated and at times even forfeited for statistical significance testing. Researchers in geography, especially those dealing with behaviour and movement, must take care to partake in this discussion and continue to develop analytical methods that can account for the heterogeneity of human spatial behaviour.

Researchers (Tellevik, 1992; Hill 1993) have used different video coding techniques to analyse and identify patterns and strategies used to explore familiar and unfamiliar environments. Most of these techniques require the recording of movement with a video camera and the isolation of different exploratory strategies by reviewing the recording as a sequence of frames. Despite a series of advances in real-time digital tracking technology (Schinazi, 2005) some researchers (Gaunet, 1996; Thinus-Blanc & Gaunet, 1999) continue to employ video based techniques in their analysis, questioning the academic value of these technological novelties. Data acquired from real-time digital tracking devices, much like video data, needs to be re-coded in order to be analyzed. The finer precision of GPS systems has allowed for tracking data to be collected and automatically coded into GIS software for analysis. This type of data collection has recently been used with considerable success in the analysis of children's activity patterns in their local environment (Mackett et al., 2006). GPS unfortunately is still not accurate enough to deal with small-scale spaces where satellite data cannot capture enough detail on the sequential propinquity of body movement and turn angles. Taken together, it seems that changes are necessary not at the *data capturing stage* but during coding and analysis. Zacharias & Schinazi (2003) have used GIS software to code and analyse the spatial behaviour in small-scale settings - the corridors of a shopping centre in Downtown Montreal.[1] However, the movement data for the most part was aggregated and

[1] Shopping centres are a good example of GPS unfriendly environment given limited strength of satellite signals.

coded as layers of polylines representing different spatial distributions and flow patterns, yet again overlooking the peculiarities of individual behaviour.

The present study will describe a technique used to examine the locomotor strategies employed by individuals who are blind or visually impaired when exploring a complex novel environment. Subjects were asked to explore a large-scale maze (45 X 30 meters), locate and remember the position of six different targets. They were then put through a series of spatial tasks and asked to make heading judgements, estimate distances and complete a cued model of the maze. The movement pattern of each subject was recorded and entered into GIS software (ArcGis) as individual polylines. The *Tracking Analyst* extension was used to view and isolate the movement patterns into specific space frames for coding. Performance in the various spatial tasks was then related to the different identified exploratory strategies.

The talk will conclude with a discussion on the complexity involved in the identification and classification of exploratory strategies that are both spontaneous and distinct. Some limitations of GIS based tracking analysis will also be presented and these will be evaluated in relation to past research and current analytical tools that allow the focus to placed on the individual.

Bibliography

Gaunet, F., & Thinus-Blanc, C. (1996). Early-blind subjects' spatial abilities in the locomotor space: Exploratory strategies and reaction-to-change performance. *Perception, 25*, 967-981.

Hill, E. W., Rieser, J. J., Hill, M. M., Hill, M., Halpin, J., & Halpin, R. (1993). How persons with visual impairments explore novel spaces: Strategies of good and poor performers. *Journal of visual impairment and blindness*, 295-301.

Mackett, R., Gong, Y., Kitazawa, K., & Paskins, J. (2006). *Where to children walk (and what do they do when they get there?).* Paper presented at the WALK21 conference, Melbourne, Australia.

Schinazi, V. (2005). *The perception and cognition of space by the blind and visually impaired. Spatial representation without vision in novel environments.* Poster presented at the Vision 2005 conference, London, UK.

Tellevik, J. M. (1992). Influence of spatial exploration patterns on cognitive mapping by blindfolded sighted persons. *Journal of visual impairment and blindness, 86*, 221-224.

Thinus-Blanc, C., & Gaunet, F. (1999). Spatial processing in animals and humans: The organizing function of representations for information gathering. In R. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (Vol. 1, pp. 427). Baltimore: The Johns Hopkins University Press.

Zacharias, J., & Schinazi, V. (2003). The impact of an entertainment retrofit on the performance of a shopping centre. *Journal of shopping centre research, 10*(1), 29-44.

Agents – Return to a computational science perspective?

Raja Sengupta, Dept. of Geography, McGill University

(A proposal for the workshop on Agent-Based Modeling of Complex Spatial Systems)

The use of the related terms "agent-based", "multi-agent", "software agent" and "intelligent agent" have witnessed significant growth in Geographic Information Science (GIScience) literature in the past decade. These terms usually refer to both *artificial life agents* that simulate human and animal behavior and to *software agents* that support human-computer interactions. While a computational science perspective does not preclude the former, Distributed Artificial Intelligence (DAI) researchers who originally coined the term and provided the semantic framework were contemplating the latter. Therefore, its extension to and usage in complexity theory research requires as a first step acknowledgement of ongoing research in the DAI community. In turn, GIScience researchers have much to offer to ongoing work on inter-agent communication languages and enforcing stronger notions of autonomy. But this will require GIScience and complexity theory agent modelers to venture into debates located deep within the unfamiliar territory of computational science.

Software agent research itself arose from the DAI community. This community was disillusioned by monolithic approaches to modeling human intelligence and believed that knowledge could be distributed into elemental components that generated emergent and intelligent behavior through interaction (Hayes-Roth and Hayes-Roth 1979, Hayes-Roth 1985, Huhns 1987, Bond and Graesser 1988). Current research on this topic within DAI builds on this basic idea using recent advances in user/software interface development (Etzioni and Weld 1994, Maes 1994), network mobility (Kotz and Gray 1999) and Internet data- and information-mining algorithms (Knapik and Johnson 1998).

Consensus amongst DAI researchers suggested that to be considered an intelligent agent, the software/computer model must possess the following four properties: (a) autonomous behavior, (b) ability to sense its environment and other agents, (c) ability to act upon its environment alone or in collaboration with others, and (d) possession of rational behavior (Woolridge and Jennings 1995, Woolridge 1999). To aid in inter-agent collaboration and communication, specific Agent Communication Languages, for example, Knowledge Query and Manipulation Language (Labrou et al. 1999), have also been developed. Additionally, researchers have pointed out that intelligent agents should not only be able to respond to, but also learn from, their environment (Maes 1994). Humanistic characteristics such as beliefs, desires, intentions (Shoham 1993), and emotions and trust (Maes 1994) also could form a part of agent behavior.

The first area of contribution by GIScience researchers can and should be in the area of inter-agent communication languages. For example, the Knowledge Query and Manipulation Language (KQML) is designed both as a messaging format and a message-handling protocol to support run-time knowledge sharing among agents. In its current format, it also includes higher level communication strategies such as contracts and negotiations. However, the ability to define spatial constructs and objects using KQML is severely limited and often non-existent. GIScience researchers, familiar with the spatial query language debate of the early 1990s (e.g., Egenhofer 1994), could capture past research on this topic to inform the expansion of KQML into its spatial equivalent (SKQML?) and in turn, enrich their own agent-modelling experience by building communicative artificial life agents.

A second area where GIScience agent researchers can assist DAI research is in implementing a "strong" notion of autonomy within artificial life agents, perhaps using virtual environments such as "Second Life" (secondlife.com) as test beds. A weak notion of autonomy suggests that agents must, in addition to being reactive, be in control of their state and persist beyond the completion of a single task (Tosic and Agha 2004). But this is true of many common software applications such as firewalls and virus scanners, and in the geospatial realm, of Internet Map Servers. A strong notion of autonomy requires the agent to have goal-directed behavior and be proactive in achieving those goals. Humans or application instantiate them but agents continue to run even after the instantiation mechanism has been terminated or is no longer present. Once instantiated, the agent must have knowledge of its goals, be in control of its actions, be able to make rational decisions in uncertain and open environments without prior knowledge about each and every situation they encounter, and require no assistance from human operators. However, in this definition, strong autonomy implies that agents have explicit spatial cognition, combined with a spatial ontology, of the virtual environments in which they operate. Spatial cognition and ontology are familiar terms to all GIScience researchers, and therefore topics on which there is much to inform the DAI community.

Bond A and Gasser L (eds) 1988 *Readings in Distributed Artificial Intelligence*. San Francisco, Morgan Kaufmann Publishers, Inc

Egenhofer MJ, 1994. Spatial SQL: A Query and Presentation Language. *IEEE Transactions on Knowledge and Data Engineering* 6(1): 86-95.

Etzioni D and Weld D 1994 A softbot-based interface to the Internet. *Communications of the ACM* 37(7): 72-6

Hayes-Roth B, Hayes-Roth F 1979 A cognitive model of planning. *Cognitive Science* 3: 275-310

Hayes-Roth B 1985 A blackboard architecture for control. *Artificial Intelligence* 26: 251-321

Hunhs M N (ed) 1987 *Distributed Artificial Intelligence*. London, Pitman Publishers

Wooldridge M and Jennings N R 1995 Intelligent agents: Theory and practice. *Knowledge Engineering Review* 10(2): 115-52

Knapik M and Johnson J 1998 *Developing Intelligent Agents for Distributed Systems.* New York, McGraw-Hill Publishers

Labrou Y, Finin T, Peng Y, 1999, Agent communication languages : the current landscape. *IEEE Intelligent Systems and their applications* 14(2) : 45-52

Tosic PT and Agha GA 2004. Towards a hierarchical taxonomy of autonomous agents. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*. Vol 4, pp. 3421- 3426.

Kotz D and Gray R S 1999 Mobile agents and the future of the Internet. *ACM Operating Systems Review* 33(3): 7-13

Maes P 1994 Agents that reduce work and information overload. *Communications of the ACM* 37(7): 30-40

Shoham Y 1993 Agent-oriented programming. *Artificial Intelligence* 60(1):51-92

Wooldridge M J 1999 Intelligent agents. In Weiss G (ed) *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, The MIT Press 27-77

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Distributed Simulation of Agent-based Models

Georgios Theodoropoulos

School of Computer Science
University of Birmingham

There has been considerable recent interest in complex systems, which involve dynamic and unpredictable interactions between large numbers of components including software, hardware devices (such as sensors), and social entities (people or collective bodies). Examples of such systems include from traditional embedded systems, to systems controlling critical infrastructures, such as defence, energy, health, transport and telecommunications, to biological systems, to business applications with decision-making capabilities, to social systems and services, such as e-government, e-learning etc. The complexity of such systems renders simulation modelling the only viable method to study their properties and analyse their emergent behaviour. Multi-agent systems (MAS) have emerged as a particularly suitable paradigm for modelling complex systems. When embedded in a real system, a MAS is itself a complex system whose properties and emergent behaviour have to be analysed via simulation. An *agent* can be viewed as a self-contained, concurrently executing thread of control that encapsulates some state and communicates with its environment and possibly other agents via some sort of message passing.

While agents offer great promise, adoption of this technology has been hampered by the limitations of current development tools and methodologies. Multi-agent systems are often extremely complex and it can be difficult to formally verify their properties. As a result, design and implementation remains largely experimental, and experimental approaches are likely to remain important for the foreseeable future. Over the last two decades, a wide range of MAS simulators and testbeds have been developed, and simulation has been applied to a wide range of MAS research and design problems, from models of complex individual agents employing sophisticated internal mechanisms to models of large scale societies of relatively simple agents which focus more on the interactions between agents. However, existing MAS simulations and simulators suffer from two key problems.

The first problem is lack of performance. The computational requirements of simulations of many multi-agent systems far exceed the capabilities of a single computer. Each agent may be a complex system in its own right (e.g., with sensing, planning, inference etc. capabilities), requiring considerable computational resources, and many agents may be required to investigate the behaviour of the system as a whole or even the behaviour of a single agent.

The second is lack of interoperability. The development of complex MAS simulation, usually requires collaborative effort from researchers with different domain knowledge and expertise, possibly at different locations. Furthermore, the effort required to develop a new simulation from scratch is considerable. There is therefore a strong incentive to reuse existing simulation components, toolkits and testbeds for a new problem. However, while many simulations have been developed, it is difficult to leverage this investment in the development of new agent simulations. Simulations developed for different agent simulators typically do not inter-operate, making it more difficult to re-use simulation components. This is particularly problematic in the case of spatial agent based models. Combining a simulation of an agent architecture developed for one simulator with a simulation of an environment developed for another typically involves re-implementation of one or both components. If the agent must be simulated in several different environments, the problem is compounded.

A solution to both these problems can be found in distributed simulation. The last decade has witnessed an explosion of interest in distributed simulation not only for speeding up simulations, but also as a strategic technology for linking simulation components of various types at multiple locations to create a common virtual environment. The culmination of this activity, has been the development of the High Level Architecture[1] (HLA), a framework for simulation reuse and interoperability developed by the US Defence Modelling and Simulation Office. Using HLA, a large-scale distributed simulation can be constructed by linking together a number of (geographically) distributed simulation components, or *federates*, into an over-all simulation, or *federation*. HLA, with minor revisions, has been adopted as an IEEE standard (IEEE 1516) and is likely to be increasingly widely adopted within the simulation community.

Distributed simulation and HLA offer an attractive potential solution to the problems of simulation and simulator reuse and simulation performance in MAS simulation. The development of HLA compliant agent simulators and simulation components would facilitate inter-operation with other simulations, allowing greater re-use of agent simulation components. In addition, the ability to distribute agent and other simulation components across multiple computers has the potential to increase the overall performance of a MAS simulation, given sufficient computational resources and favourable simulation characteristics. However most of the work in this area to date has employed various ad-hoc approaches to parallel simulation, e.g., distributing the agents over a network of processors interacting via some communication protocol, and has yielded relatively poor performance. MAS models present particular challenges for distributed simulation. An example key problem in the distributed simulation of spatial agent-based models is the efficient distribution of the agents' environment, namely the part of the world (or computational system) 'inhabited' by the agent. In simulations of situated MAS, the environment is represented by a large shared state space, which may

[1] https://www.dmso.mil/public/transition/hla/

be accessed by any of the agents frequently and in dynamic, non-deterministic patterns. It is therefore difficult to determine an appropriate simulation topology *a priori*. Encapsulating the shared state in a single process (e.g. via some centralised scheme) introduces a bottleneck, while distributing it all across the distributed resources (decentralised, event driven scheme) will typically result in frequent all-to-all communication and broadcasting.

While the HLA enables interoperability and the construction of large-scale distributed simulations using existing and possibly distributed simulation components, it does not provide support for collaborative development or configuration of simulation applications, nor does it provide any mechanism for managing the resources where the simulation is being executed. The emergence of Grid technologies provides exciting new opportunities for large scale distributed simulation of agent-based models, enabling collaboration and the use of distributed computing resources, while also facilitating access to geographically distributed data sets.

How should large-scale environmental data sets be distributed to enable efficient access by the agents in a distributed simulation? How can we develop efficient infrastructures to support load management, synchronisation, and query routing in large scale MAS simulations? How can we support collaborative model development and automated composition of simulation components? Is HLA the answer to interoperability? The workshop can discuss some of these challenges, which are at the heart of making agent-based simulation feasible at a scale necessary to master the challenges ahead of us.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

A Personal Perspective on Agent-Based Modelling of Complex Spatial Systems

Andy Turner
http://www.geog.leeds.ac.uk/people/a.turner/

Preparation for the NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems 14th to 16th April 2007, Santa Barbara USA.

Introduction

As a computational geographer, I attempt to model space-time-attribute interactions at a human scale and for a set of applied uses. My focus is contemporary and in 2006 began a PhD topic on Agent Based Models (ABM) of daily activity. In this I plan to develop a model of the UK operating at the individual person level and run some simulations at as detailed a spatial and temporal resolution as I can manage using available computational resources.

I am generally interested in the evolution of socio-economics, in particular, resource discovery and exploitation from the agrarian revolution through the development of nation states and the industrial revolution and into the on-going information revolution.

On ABM

Agent Based Modelling is object orientated modelling with objects that interact and that may change behaviour or state as a consequence of this interaction. In ABM, a sometimes simple and sometimes complex *environment* object is often not regarded as an agent, but distinguished as a special object. The environment is the collection of all agents and a spatio-temporal framework in which agents operate. The environment is often not regarded as an agent although it has characteristics that change in response to agents actions. These changes are often fed back into changes in agent behaviour. In considering the environment as a framework or arena of action there are spatial and temporal origins, extents, scales and resolutions. What are the effects of arbitrary choices of these?

On Games, Education and Decision Support

There are some fantastic agent based models used in the entertainments industry, particularly in computer games. Virtual worlds are in genesis using software and models produced by game developers and models and simulations modified and played out by user communities. The games can end up incorporating detailed knowledge about sequences and timings of historical events; socio-economic, political and technological. How can we harness this and collaborate to generate planning and educational tools for operational social government?

Social shaping and the interaction between virtual environments and the movement of people in society although an interesting feedback, is perhaps a topic for another workshop another time.

Topics of relevance

- Data

    - Archives
    - Provenance
- Technology
    - High Performance Computing
    - Distributed Computing
    - Open Source Software
- Uncertainty
- Correlation of observed and estimated characteristics
- The potential of subjective models that cannot be validated
- Visualisation
- Linking models
- Scale and Resolution

On the workshop

I am looking forward to this event. I enjoy establishing and developing collaborations. I'm looking forward to reading about you and more about the subject in preparation.

As this is in part an NCeSS workshop, there are collaboration tools we can use. In particular, a wiki (http://www.ncess.ac.uk/support/wiki/). Anyone can get an account on this and I aim to use it to develop content as part of our collaboration. I have started a page for this workshop, it is nothing more than a stub at the time of writing. It is linked from the main page in the Agenda Setting Workshop list as Agent Based Modeling of Complex Spatial Systems (http://www.ncess.ac.uk/support/wiki/index.php/ABMofComplexSpatialSystems).

APPLICATION FOR THE SPECIALIST WORKSHOP ON AGENT-BASED MODELING OF COMPLEX SPATIAL SYSTEMS

Stephan Winter
Department of Geomatics, The University of Melbourne, VIC 3010, Australia
winter@unimelb.edu.au

1. PERSONAL PERSPECTIVE ON THE TOPIC OF THE WORKSHOP

Over the past few years I recognized in my research that the community of agent-based modelers of spatial systems on the one hand, and the community interested in complex dynamic systems on the other developed more or less independently from each other. And I became interested in studying potential relationships between the realms of the communities for mutual benefits and discovery of new knowledge. The complex dynamic systems I am working on are forms of social networks, more particularly ad-hoc social networks—agents that are near to each other and hence can interact. The latter brings in a geographic perspective, which can become more relevant if the matter of interaction is solving spatial problems, e.g., wayfinding, or spatiotemporal problems, e.g., transportation planning.

1.1 An Example

Research on geosensor networks is typically concerned with the efficient extraction of information of sensor observations, hence, looking into hardware, protocols, routing of messages, and data aggregation, acknowledging that geosensor network nodes are mobile and always aware of their location. My research focus is different in three respects.

- First of all, its focus is on movement of the nodes, not of information. The investigated geosensor network consists of nodes that have individual travel intentions. If two nodes meet, one of them can ride piggy-back on the other one for reasons such as saving fuel or traveling faster, depending on the abilities of the two nodes. We call this behavior ride sharing, and are interested in all forms of travel optimization.
- Secondly, this geosensor network allows for different types of nodes. In applications, one will distinguish transportation clients—nodes that can travel piggy-back—from transportation hosts—nodes that offer piggy-back traveling. Some clients may be able to move only with a host—think of parcels in a freight application—, while other clients may be able to move independently, for example pedestrians in an urban transport application. There may also be hosts that offer piggy-back traveling not only to clients, but also to other hosts, as for example a ferry to vehicles. Even immobile nodes can be thought of as part of the geosensor network, participating in the peer-to-peer communication—think of bus stops in a public transportation application, mediating between buses and pedestrians. With all their individuality, nodes can be conceptualized as agents.
- Thirdly, special challenges arise from the typical communication constraints of geosensor networks: scarce resources in terms of battery and bandwidth, and of a fragile communication network topology due to node mobility. Nodes have to communicate with each other to gather current transportation network knowledge for trip planning. Being restricted to local communication means that nodes can only gather local transportation network knowledge, and hence, find sub-optimal trips.

The interesting research questions in this context are on the nodes' communication strategies, on their (optimal) trip planning strategies, on competition, on trust and reputation in peer-to-peer systems, on the general behavior of large transportation geosensor networks considering the autonomy and intentionality of the nodes, or on the potential for heterogeneous network architectures (wired/wireless), to name just a few. While these questions give a hint of the challenges, I will illuminate the questions by a concrete realization: a shared ride system for persons traveling by multiple modes in the city.

The envisioned peer-to-peer shared ride system enables pedestrians to negotiate in an ad-hoc manner for ride sharing with diverse vehicles in urban traffic, such as private cars, buses, trains, or taxi cabs. In such a system pedestrians are the clients, and vehicles are the hosts. Finding rides in an ad-hoc manner is accomplished by local negotiation between these agents via radio-based communication. For this purpose, clients collect local real-time transportation network information. Based on this information and their preferences for various optimization criteria, such as travel fees or travel time, they are able to select hosts that offer optimal trips with respect to their limited information. The selected hosts are booked then, and the ride can take place. Since the selection was made based on local knowledge, and other opportunities could come along, the client revises its travel plans regularly.

In previous research we have investigated the clients' ability to make trip plans from local transportation network knowledge. Optimizing for travel time in this case, the results show that local communication is both efficient and effective. It is efficient because it leads to less communication messages than for complete current transportation network knowledge, and it is effective because it leads to travel times comparable to the travel times of trips computed from complete current knowledge. This investigation was realized by simulation of a peer-to-peer shared ride system, with an immobile and inflexible client agent and homogeneous host agents. We were concerned whether the simulation design was elaborate enough to reflect sufficiently the behavior of a system deployed in the real world. For this reason, the simulation was extended, introducing diverse types of client and host agents with different behavior and capabilities, including deterministic mobility models. It turned out that with every step to more complexity, local communication is always both efficient and effective compared to other communication strategies, even if the trips themselves may change significantly. For details on this work see the references in Section 2.2.

1.2 Some Observations from the Example

Current multi-agent systems, such as Swarm, Repast, or Ptolemy, are lacking spatial awareness. Developing ad-hoc alternative systems seems not a viable solution.

- On the agenda: engage in public domain projects and contribute spatial abilities to multi-agent systems, such as spatial data and knowledge representation, route planning and following, or a spatial communication layer.

Multi-agent systems are perfect tools to simulate mobile sensor networks. Simulations with multi-agent systems, in particular in the area of mobile sensor networks, use frequently simple mobility models such as random walking models.

- On the agenda: engage in the development of mobility models better reflecting goal-oriented behavior of typically repetitive patterns.

Multi-agent systems that have other mobility models (micro-simulation models) are designed for realistic traffic modeling (Nagel 2001; Torrens 2004) or modeling of dynamic urban processes (Batty et al. 2003; Benenson and Torrens 2004).

- On the agenda: engage in models of cooperation and communication, since real-world agents will more and more being ubiquitously connected by some communication technology and do interact.

2. RESUME

I am Senior Lecturer at the Department of Geomatics, The University of Melbourne, Australia. Previously I held positions at the University of Bonn, Germany (1990-97, PhD in 1997), and the Technical University Vienna, Austria (1997-2003, habilitation in 2001).

My current research focuses on spatial information theory with a specific interest in spatial cognition, wayfinding and navigation. I also look into conceptual route planning, network analysis, and, recently, mobile ad-hoc geosensor networks and their application in real-time and dynamic route planning. My preferred research approach is simulating complex dynamic systems by modeling autonomous agents, and analyzing the decision strategies of the agents with respect to optimality or efficiency. I am also member of the virtual sensor network laboratory of the University of Melbourne, providing a testbed environment for ideas.

Currently I am leading a CRCSI project on adapting route information for different user groups, and recently I have won an ARC Linkage project on cognitive ergonomic wayfinding directions. Another project, on peer-to-peer shared ride systems, runs without major external funding. Two years ago I started a collaboration with Silvia Nittel, head of the Geosensor Network Lab at the University of Maine, and looked into peer-to-peer shared ride trip planning with mobile location-aware sensor networks (Winter and Nittel, 2006). These papers explore whether this novel approach to improve urban mobility is feasible, which communication strategies and protocols are required and efficient, and what the next research questions are. In the next paper, with Martin Raubal, at that time University of Muenster, I proposed algorithms from time geography to identify relevant data for shared ride planning, by this way reducing peer-to-peer communication to a minimum (Winter and Raubal, 2006).

Previous research concerned topological relations in presence of location uncertainty. Working in this area I applied a model of raster-vector-unification for topological analysis, and reused it later for formal specifications for interoperability. Parts of these research results were included in the OpenGIS specifications. During this time I participated in several research projects on interoperability, before I won European and national research projects in the area of navigation and tourist information.

My publication record notes more than 100 publications over the last 10 years, 42 of them full paper reviewed. Among these papers is the most cited GIScience paper (according to Google Scholar, September 2005), written together with Martin Raubal. Courses I am teaching deal with spatial data management and analysis, webmapping and interoperability, and navigation services.

In 2007 I will chair the Eighth International Conference on Spatial Information Theory (COSIT’07), together with Benjamin Kuipers, Texas. In conjunction with COSIT’07, I will also chair the first international workshop on social space and geographic space, looking into social networks and their geographic conditions. This workshop will be sponsored by the ARC Network of Spatially Integrated Social Sciences (ARCNSISS), of which I am founding member. I chaired already an ARCNSISS workshop on trust and reputation in ad-hoc local communities. In 2000, I chaired the EuroConference on Ontology and Epistemology for Spatial Data Standards in France, 2000. I am co-chair of the ISPRS WG II/6: Geo-spatiotemporal semantics and interoperability, and chaired the *Working Group in Interoperability* in the Association of GI Laboratories in Europe until 2003.

2.1 My ten career-best publications

Claramunt, C.; **Winter**, S., forthcoming: Structural Salience of Elements of the City. Environment and Planning B, accepted for publication in October 2006.

Raubal, M.; **Winter**, S.; Tessmann, S.; Gaisbauer, C., forthcoming: Time Geography for Intelligent Ad-Hoc Shared-Ride Trip Planning. International Journal of Photogrammetry and Remote Sensing (Special Issue on From Sensors to Systems: Advances in Distributed Geoinformatics), accepted for publication in December 2006.

Winter, S.; Nittel, S., 2006: Ad-Hoc Shared-Ride Trip Planning by Mobile Geosensor Networks. International Journal of Geographical Information Science, 20 (8): 899-916.

Klippel, A.; **Winter**, S., 2005: Structural Salience of Landmarks for Route Directions. In: Cohn, A.G.; Mark, D.M. (Eds.), Spatial Information Theory. Lecture Notes in Computer Science, Vol. 3693. Springer, Berlin, pp. 347-362.

Nothegger, C.; **Winter**, S.; Raubal, M., 2004: Computation of the Salience of Features. Spatial Cognition and Computation, 4 (2): 113-136.

Winter, S.; Nittel, S., 2003: Formal Information Modeling for Standardisation in the Spatial Domain. International Journal of Geographical Information Science, 17(8):721-741.

Raubal, M.; **Winter**, S., 2002: Enriching Wayfinding Instructions with Local Landmarks. In: Egenhofer, Max J.; Mark, David M. (Eds.), Geographic Information Science. Lecture Notes in Computer Science, Vol. 2478. Springer, Berlin, pp. 243-259.

Winter, S., 2002: Modeling Costs of Turns in Route Planning. GeoInformatica, 6(4): 345-360.

Winter, S., 2000: Uncertain Topological Relations between Imprecise Regions. International Journal of Geographical Information Science, 14(5): 411-430.

Winter, S.; Frank, A.U., 2000: Topology in Raster and Vector Representation. GeoInformatica, 4(1): 35-65.

2.2 Any papers already published by me in the area of the workshop contribution

Winter, S.; Nittel, S., 2006: Ad-Hoc Shared-Ride Trip Planning by Mobile Geosensor Networks. International Journal of Geographical Information Science, 20 (8): 899-916.

Winter, S.; Raubal, M., 2006: Time Geography for Ad-Hoc Shared-Ride Trip Planning. In: Aberer, K.; Hara, T.; Joshi, A. (Eds.), 7th International Conference on Mobile Data Management (MDM'06). IEEE Computer Society, Nara, Japan, Paper 6.

Extended version accepted as:

Raubal, M.; Winter, S.; Tessmann, S.; Gaisbauer, C.: Time Geography for Intelligent Ad-Hoc Shared-Ride Trip Planning. International Journal of Photogrammetry and Remote Sensing (Special Issue on From Sensors to Systems: Advances in Distributed Geoinformatics), forthcoming.

Wu, Y.-H.; Guan, L.-J.; Winter, S., 2007: Peer-to-Peer Shared Ride Systems. In: Nittel, S.; Stefanidis, A.; Labrinidis, A. (Eds.), 2nd International Conference on GeoSensor Networks. University of Maine, Boston, MA, pp. 27-38.

Revised version accepted for:

Nittel, S.; Labrinidis, A.; Stefanidis, A. (Eds.), GeoSensor Networks 2006. Lecture Notes in Computer Science, Springer, Berlin, forthcoming.

3. References

Batty, M.; Desyllas, J.; Duxbury, E., 2003: The Discrete Dynamics of Small-Scale Spatial Events: Agent-Based Models of Mobility in Carnivals and Street Parades. International Journal of Geographical Information Science, 17 (7): 673-697.

Benenson, I.; Torrens, P.M., 2004: Geosimulation: Automata-based Modeling of Urban Phenomena. John Wiley & Sons, Chichester, UK, 336 pp.

Nagel, K., 2001: Multi-modal traffic in TRANSIMS. In: Schreckenberg, M.; Sharma, S.D. (Eds.), Pedestrian and Evacuation Dynamics, Springer, Berlin, pp. 161-172.

Torrens, P.M., 2004: Geosimulation, Automata, and Traffic Modelling. In: Hensher, D.A. et al. (Eds.), Handbook of Transport Geography and Spatial Systems. Handbooks in Transport, 5. Elsevier, Amsterdam, pp. 549-564.

Mobile P2P Databases[1]

Ouri Wolfson
Department of Computer Science
University of Illinois at Chicago

A mobile peer-to-peer (P2P) database is a complex spatial-temporal system, in which information is stored in the peers of a mobile P2P network. The mobile peers communicate with each other via short range wireless protocols, such as IEEE 802.11, Bluetooth, Zigbee, or Ultra Wide Band (UWB). These protocols provide broadband (typically tens of Mbps) but short-range (typically 10-100 meters) wireless communication. On each mobile peer there is a local database that stores and manages a collection of data items, or reports. A report is a set of values sensed or entered by the user at a particular time, or otherwise obtained by a mobile peer. Often a report describes a physical resource such as an available parking slot. All the local databases maintained by the mobile peers form the mobile P2P database. The peers communicate reports and queries to neighbors, and these propagate by transitive multi-hop transmissions.

Mobile P2P databases enable matchmaking or resource discovery services in many application domains, including social networks, transportation, mobile electronic commerce, emergency response, and homeland security.

Traditionally search databases have been implemented by a centralized architecture. Google is preeminent example of such architecture. However, mobile P2P databases have several advantages over centralized ones, including higher reliability, better privacy and performance, lower cost, and independence of a fixed infrastructure. Their disadvantage is that they do not guarantee answer completeness.

The concept of mobile P2P database is proposed for searching local information, particularly information of a temporary nature, i.e. valid for a short duration of time [1].
There are two main paradigms for answering queries in mobile P2P databases. One is pulling reports by sending queries on search missions in the network, and the other is pushing the reports to the right queries. Combination approaches seem most promising.

There are many research challenges in mobile P2P databases:

1. **Prolong network lifetime:** Currently, some approaches e.g. ranking and cluster-based-methods, are proposed to prolong the lifetime of sensor networks, mobile ad hoc networks, and mobile P2P databases. The future research question is how to employ the redundancy of networks and the density of peers in order to maximally extend the network lifetime.

[1] Research supported by NASA Award NNA06AA25C, and NSF Grants OII-0611017, DGE-0549489, 0513736, and 0326284.

2. **Sparse networks**: Currently, the performance of many algorithms and systems heavily depends on the density of peers in mobile P2P networks. They do not perform very well if the network is sparse. Therefore, how to design and develop mobile P2P databases for sparse networks is an important and difficult challenge. Recent work that heads in this direction includes Delay Tolerant Networks, store and forward flooding, and mobile peers whose sole function is to provide connectivity.
3. **Rapid topology changes:** Highly mobile peers pose problems, e.g. how to efficiently disseminate queries and answers, and how to reconfigure rapidly when the topology of networks changes frequently. Stateless approaches seem most suitable to address these problems.
4. **Emergent global behavior from local knowledge:** Mobile P2P databases can be treated as a special type of distributed system. Each peer maintains a local database and all the local databases form the virtual mobile P2P database. Therefore, peers can only use the local knowledge to predict or affect the global behavior of the whole mobile P2P database. The future research direction is how to employ the local knowledge and propose adaptive local algorithms to direct or affect the global behavior of mobile P2P databases.
5. **(Self-) Localization techniques:** Location-based approaches are increasingly popular and necessary, and location information of peers is useful for efficiently storing and managing information. However, self-localization techniques are still not efficient and effective enough due to the limitation of peers or localization techniques. For example, GPS is not available indoors and the accuracy of GPS is not enough for some mobile P2P databases. Therefore, efficient and effective self-localization technique for mobile P2P databases is an important research direction.
6. **Mathematical modeling of data dissemination:** Many query processing and data dissemination algorithms may benefit from a mathematical model of data propagation. For example, a formula giving the number n of mobile peers having a report that was generated at time t at location l would be very useful in ranking of such a report. The number n is a function of the density of mobile peers, motion speed, bandwidth and memory availability at the peers, memory management, etc. Related work done in epidemiology about the spread of infectious diseases would be a good starting point for this research. Results in random graphs are also applicable.

Other important research directions include incentives for broker participation in query processing, transactions/atomicity/recovery issues in databases distributed over mobile peers, answering specialized queries that are amenable to specific optimization, and integration with the fixed infrastructure.

[1] B. Xu and O. Wolfson, "Data Management in Mobile P2P Networks", *Springer Verlag Lecture Notes in Computer Science, Proc. of the 2nd International Workshop on Databases, Information Systems, and P2P Computing (DBISP2P'04)*, Toronto, Canada, Aug 2004.

[2] M. Motani, V. Srinivasan, and P. Nuggehalli, "PeopleNet: Engineering a Wireless Virtual Social Network," *International Conference Mobile Computing and Networking (MobiCom'05)*, Aug 2005.

Personal Perspective on the Workshop on Agent-Based Modeling of Complex Spatial Systems

Michael Worboys

In 1988, Mark Weiser, a chief technologist at Xerox's Palo Alto Research Center, introduced the term ubiquitous computing, and heralded "the age of calm technology, when technology recedes into the background of our lives." This vision is now fast becoming a reality. Computing will soon be embedded everywhere in the fabric of our lives: in our bodies, phones, homes, and the environments in which we live. Physically located sensors will be collecting data from a multiplicity of diverse sources using a variety of sensor technologies. Sensors will be mobile and the sensed phenomena are dynamic. Computing is expanding from people's desktops to wherever people are, making it an integral part of people's lives. At the same time, computing is also spreading to everywhere independent of the presence of people, leading to ubiquitous computing environments. Such anytime-anywhere computing revolutionizes the way people will use and interact with computers in the future outside their traditional office or home environments. Example scenarios for the use of such technology include emergency response, smart transportation systems, and real-time environmental monitoring.

Complex spatial models also manifest in dynamic field applications. Natural hazards, such as floods, fires, earthquakes, and volcanic eruptions provide significant threats to health and the economy. Pollution, such as that caused by carbon dioxide emissions, poses similar problems. Some of these phenomena are infrequent, but require intensive observation over short periods. Others are chronic and require different spatio-temporal patterns of monitoring. What many such phenomena have in common is the creation and dispersal of *dynamic fields*, whether of a gas, a level of temperature, of seismic activity, or water level. Ubiquitous spatial computing technology, in the form of sensor networks, also provides support for these applications.

Does geographic information science have a role to play in these new technological developments? My strong belief is that spatial and spatio-temporal issues lie at the very heart of ubiquitous computing. Unlike the virtual reality paradigm, ubiquitous computing is computing in the real, physical world. Event-based and agent-based models are key.

The perspective described here concerns the information-theoretic foundations upon which useful explanatory and predictive models of dynamic geographic phenomena sensed in ubiquitous computational environments can be based. We see a development of these foundations, from sequences of temporal snapshots, through object life histories, to event chronicles. A crucial ontological distinction is drawn between "things in the world" and "happenings in the world"; that is, between continuant and occurrent entities. Up to now, most research has focused on representing the evolution through time of geographic continuant entities, whether objects or fields. This paper argues that occurrents should be upgraded to an equal status with things in models of dynamic phenomena.

We propose several approaches to the resulting modeling questions. The underlying framework is provided by a formal approach to dynamic topology. The entities under investigation here are topological occurrents, such as splitting, merging, hole development, and other changes to connectivity relationships. This framework provides the basis for developing an approach to dynamic object-based and field-based views.

It is possible to develop a pure event-oriented theory of space and time, where spatio-temporal locations and agents situated in the space are all viewed as processes. We can then apply algebraic methods, such as the various process and event calculi developed by computer scientists. However, so far, it has been difficult to demonstrate the scalability of such an approach: the complexities of formal representation of even simple real-world events become quite forbidding. Nevertheless, we believe that a modular approach to such representations provides a way forward in this area, in a similar way that object encapsulation and modularisation provided a way forward in object-oriented approaches. The advantage of such an approach is that the formalism has a great deal more power, both in terms of representing complex processes and in reasoning about them.

Another area of interest in complex spatial systems is the modeling of sensor networks in dynamic fields. We have been working on an approach to information management that uses a qualitative approach to identifying and tracking continuous environmental phenomena such as toxic clouds or oil spills. One of the key elements of our approach is the modeling of the information management infrastructure in the underlying sensor network by the use of combinatorial maps. A sensor node locally stores information about the nodes in its communication neighborhood as a set of "darts". These darts are ordered based on their spatial direction in a cyclic order around the sensor. We are researching the underlying theory and algorithms associated with this continuously adapting information management infrastructure. These algorithms assume no centralized control and are highly distributed. Key components of this research are:

- Detection of qualitative changes to the dynamic field by a sensor or group of sensors.
- Appropriate response to these changes in terms of local reconfigurations of the network and data routing.

The common theme throughout this work is models complex, dynamic phenomena, with emphasis on models that are:

- Agent-based, focusing on the processes that agents can perform, and their properties.
- Event-based, focusing on the occurrent entities in the phenomena, and their relationships.
- Distributed, not admitting centralized control, and therefore concerned with emergent properties.
- Rich, allowing a high degree of representational and reasoning power.

NSF/ESRC Agenda Setting Workshop on Agent-Based Modeling of Complex Spatial Systems: April 14-16, 2007

Belinda Wu, Leeds

A Hybrid Approach for Spatial MSM

Computer simulations/models have now become more important in modelling complex systems including the social systems. Modern policy problems often require disaggregate information with great details. IBM (Individual Based Model) models the system at the individual level. IBM can provide individual characteristics to assist decision making in contrast to the traditional models where individual characteristics are often blurred or even disappeared.

MSM (Microsimulation Model) and ABM (Agent Based Model) are the two important approaches in IBM. MSM is an extensively applied and well proven approach in social modelling. Especially in the public policy domains, its application has ranged from tax-benefit, pension, health to transport policies (Redmond *et al*. 1998; Sutherland, 2001; Curry, 1996; Morrison, 2003; PTV AG, 2000). Spatial MSM simulates virtual populations in given geographical areas (Ballas et al, 2005) so that local contexts can be taken into account when studying the characteristics of these populations and analysing the policy impacts (Birkin and Clarke, 1995; Clarke, 1996).

Although limitations such as data and computation requirements have been greatly improved nowadays, two criticism against MSM remain to be addressed: MSM are less strong in behaviour modelling and most MSM only models one-direction interactions: the impact of the policy on the individuals, but not the impact of individuals on the policy (Krupp, 1986; Williamson, 1999; Citro and Hanushek, 1991; O'Donoghue, 2001; Gilbert and Troitzsch, 2005).

ABM can provide the capability for behaviour modelling. It allows us to study the interactions between the policy and population at both macro and micro levels, as well as in both directions. Agent based social simulation can provide insight into the structure and effects of policies and can assist in understanding and modifying behaviour and interaction patterns (Luck et al., 2003). However, despite the usefulness of the ABM as described in previous discussion, being a relatively new technology, sometimes it is felt that it can benefit from more refined and well-established theories and concepts of other approaches (Gilbert and Troitzsch, 1999; Conte *et al.*, 1998). Such features make the MSM and ABM naturally complement each other.

MoSeS proposes a hybrid modelling approach that brings the strength of the MSM and ABM together based on four considerations:

- MSM and ABM complement each other;
- Geography provides a bridge to link the MSM and ABM;
- Previous attempts of hybrid approaches have resulted in fruitful outcomes;

- This hybrid approach may provide a new angle to view classical problems (Boman and Holm, 2004).

References:

Ballas, D.; Clarke, G.; Dorling, D.; Eyre, H.; Thomas, B. and Rossiter, D. (2005) SimBritian: A spatial microsimulation approach to population dynamics, in Population, Space and Place 11, 13-34.

Birkin, M. and Clarke, G. (1995) Using microsimulation methods to synthesize census data', in S. Openshaw (ed.) *Census Users' Handbook*. London: GeoInformation International, pp. 363–87

Boman, M. and Holm, E. (2004) Multi-agent systems, time geography, and microsimulations, in Olsson, M-O. and Sjöstedt, G. (eds.) Systems Approaches and their Application, 95--118, Kluwer Academic.

Citro, C. and Hanushek, E. (1991) Improving information for social policy decisions: The uses of microsimulation modelling, vol. 1, Review and recommendations. Washington, DC: National Academy Press.

Clarke, G.P. (1996) 'Microsimulation: an introduction', in G.P. Clarke (ed.) Microsimulation for Urban and Regional Policy Analysis. London: Pion, pp. 1–9

Conte, R. and Gilbert, N. (1995) Introduction, in Gilbert and Conte (eds.) Artificial Societies: The Computer Simulation of Social Life, London.

Curry, C. (1996) PENSIM: A Dynamic Simulation Model of Pensioners' Income, Government Economic Service Working Paper No. 129, Analytical Services Division, Department of Social Security, London

Gilbert, N. and Troitzsch, K. (1999) Simulation for the Social Scientist, Open University Press.

Krupp, H. (1986) Potential and limitations of microsimulation models, in Orcutt *et al.* (eds) Microanalytic simulation models to support social and financial policy, North-Holland, Amsterdam

Luck, M. and McBurney, P. and Preist, C. (2003) Agent Technology: Enabling Next Generation Computing (A Roadmap for Agent Based Computing). AgentLink.

Morrison, R.J. (2003) Making Pensions out of Nothing at All, The International microsimulation Conference on Population, Ageing and Health: Modelling our Future.

Moss, S. (2000) Messy Systems - The Target for Multi Agent Based Simulation, In Moss, S. and Davidsson, P. (Eds.) Multi Agent Based Simulation, Springer Verlag LNCS series, Vol. 1979.

O'Donoghue, C (2001) Dynamic microsimulation: A methodological survey. Brazilian Electronic Journal of Economics 4 (2).

PTV AG (2000), VISSIM User manual - Version 3.00, Karlsruhe, Germany

Redmond, G.; Sutherland, H. and Moira Wilson, M. (1998) The Arithmetic of Tax and Social Security Reform: A User's Guide to microsimulation Methods and Analysis, Cambridge: Cambridge University Press

Sutherland, H. (2001) EUROMOD: An Integrated European Benefit-tax Model. Final Report, EUROMOD Working Paper No EM9/01, DAE University of Cambridge, Cambridge.

Williamson, P. (1999) Microsimulation: An idea whose time has come? the 39th European Congress of the European Regional Science Association, August 23-27.

Temporal GIS for Agent-Based Modeling of Complex Spatial Systems
By May Yuan, University of Oklahoma

Agent-based modeling has become one of the key computational approaches to simulate collective outcomes out of individual decisions in complex spatial systems. Much effort has been devoted to identifying, formulating, and experimenting with rules of local behavior for discovery of emergent, self-organizing global patterns. With emphases on computation, agent-based modeling mostly operates on cell-, lattice-, or network-based data structures (Batty 2005; Andersson et al. 2006; Andersson et al. 2006; Bithell and Macmillan 2007). While agent-based modeling aims at discerning higher orders from complex disintegrated actions, it is limited by these confined data structures that restrict neighborhood geometry and possible locations, spatial interaction structures, and local spatial scale on actions and interactions. On one hand, agent-based modeling attempts to capture spatial complexity, but on the other hand, spatial data structures used for the modeling approach inevitably over-simplify the complex nature of geographic space. In this position paper, my premise is posited upon the need for temporal GIS representation of complex properties that manifest spatiotemporal presence of geographic dynamics.

There are at least three dimensions of complexity in geographic dynamics (Goodchild et al. 2007): changes to geometry, changes to location (movement), and changes to internal structure. Based on the three dimensions, agent-based modeling can be considered a means to simulate individual movements in order to examine aggregated changes to geometry and internal structure. Therefore, agent-based modeling adds another layer of complexity: the inherited hierarchical nature of geographic dynamics that propagates from individuals to the aggregated whole. Furthermore, individuals and aggregates are relative concepts. An aggregate at one level may be considered an individual in a higher level. For example, a residential district is an aggregate of houses, but a district can also be considered an individual that aggregates to a community. Multiple levels of aggregation over geographic semantics, space, and time are outcomes of dynamics that operate at and across different spatiotemporal scales. A hierarchy of geographic dynamics also suggests the potential for distinct rules for actions at different levels in the hierarchy. Group behavior and psychology likely depart from individuals'. If temporal GIS representation can capture the intricate, multi-level and complex structure of geographic dynamics, the temporal GIS can empower agent-based modeling in two significant ways discussed below.

First, rules applicable to different levels of geographic dynamics can be incorporated into agent-based modeling. For example, individual drivers can be regarded as fine-grain agents, and comparably convoys of vehicles like coarse-grain agents. Fine-grain agents apply different rules of actions than coarse-grain agents, even though coarse-grain agents may be aggregates of correspondent fine-grain agents. GIS data representing a lower level of geographic dynamics (such as traffic signs and traffic counts) provide the basis for agent behavior at a finer grain. GIS data representing a higher level of geographic dynamics (such as traffic flows and highway types) offer the condition for agents of a coarser grain. Emergent patterns can then be observed at multiple levels of detail. Simon (1973) argued that any complex system in the world must be hierarchical, or otherwise we would have no way to acquire it. He further elaborated on the importance of hierarchical structures to the sustainability of a complex system, for only hierarchies can evolve efficiently and successfully in a consistently changing world. While reality may or may not be hierarchical, hierarchical structures facilitate observations and understanding (Allen and Starr 1982). Agent-based modeling needs to incorporate the hierarchical nature of geographic dynamics, and temporal GIS needs to support the necessary data in forms that enable the simulation of agent actions within and across levels of geographic dynamics.

Second, temporal GIS can provide empirical support if the results from agent-based modeling can be stored into a spatiotemporal database for query, retrieval, and analysis. The empirical support will allow for comparison of model output and observations and comparison between modeling results from different scenarios or rule sets. Such comparisons can be change-based or development-based. Emphases on change center on the differences in distributions and patterns in space and time. Examples include how spatial distributions of pedestrians change over time, and how emergent patterns (shape and topology) differ based on different sets of behavior rules. Development-based comparisons focus on the evolution of individual agents or groups of agents. For instance, an agent adapts to environmental conditions at finer and coarser grains, and a forest may diminish and become more fragmented over time. When the results can be stored in a temporal GIS database, algorithms can be developed to support queries that seek similarity from empirical observations or from model runs with different rules. A temporal GIS representation framework that combines field- and object-based models to capture precipitation change and rainstorm development (Yuan 2001; McIntosh and Yuan 2005; McIntosh and Yuan 2005) can be revised to enable such empirical support to agent-based modeling.

An integration of agent-based modeling (ABM) and temporal GIS (TGIS) data modeling offers both theoretical and empirical improvements to understand spatial complex systems. Agent-based modeling can effectively represent the distributed nature of actions and reactions at the individual levels and transcend the individual, local decisions to identifiable patterns at a higher level in a complex spatial system. Temporal GIS data modeling have advanced to represent geographic complexity and dynamics and organize spatiotemporal data according to processes measured/recorded by these data. Therefore, the ABM-TGIS integration promises novel approaches to the study of spatial complex systems.

Andersson, C., K. Frenken and A. Hellervik (2006). A complex network approach to urban growth. *Environment and Planning A* 38(10), 1941-1964.
Andersson, C., K. Frenken and A. Hellervik (2006). A complex network approach to urban growth. *Environment and Planning A* 38(10), 1941-1964.
Batty, M. (2005). Agents, cells, and cities: New representational models for simulating multiscale urban dynamics. *Environment and Planning A* 37(8), 1373-1394.

Bithell, M. and W. D. Macmillan (2007). Escape from the cell: Spatially explicit modeling with and without grids. *Ecological Modelling* 200, 59-78.

Goodchild, F. M., M. Yuan and T. Cova (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science* 21(3), 239-260.

McIntosh, J. and M. Yuan (2005). Assessing similarity of geographic processes and events. *Transactions in GIS* 9(2), 223-245.

McIntosh, J. and M. Yuan (2005). A framework to enhance semantic flexibility for analysis of distributed phenomena. *International Journal of Geographic Information Science* 19(10), 999-1018.

Yuan, M. (2001). Representing Complex Geographic Phenomena with both Object- and Field-like Properties. *Cartography and Geographic Information Science* 28(2), 83-96.